

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/092000 A2

(51) International Patent Classification⁷: **A61K**
(21) International Application Number: PCT/US02/14877
(22) International Filing Date: 13 May 2002 (13.05.2002)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
60/290,071 11 May 2001 (11.05.2001) US
60/291,311 17 May 2001 (17.05.2001) US
60/353,058 1 February 2002 (01.02.2002) US
60/361,293 4 March 2002 (04.03.2002) US

[US/US]; 11 Oliver Lane, Hopkinton, MA 01748-3108 (US). **ANISOWICZ, Anthony** [US/US]; 50 Upham Street, West Newton, MA 02465 (US). **GRAHAM, James, R.** [US/US]; 40 Peirce Street, Arlington, MA 02476 (US). **MORALES, Arturo** [US/US]; 230 Cedar Ave., Arlington, MA 02476 (US). **YAWORSKY, Paul, J.** [US/US]; 13 Hobart Lane, Rockland, MA 02370 (US). **LIU, Wei** [CN/US]; 73 Blackmer Road, Sudbury, MA 01776 (US).

(74) Agents: **REA, Teresa, Stanek et al.**; BURNS, DOANE, SWECKER & MATHIS, L.L.P., P.O. Box 1404, Alexandria, VA 22313-1404 (US).

(71) Applicants (for all designated States except US): **GENOME THERAPEUTICS CORPORATION** [US/US]; 100 Beaver Street, Waltham, MA 02453 (US). **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07928 (US).

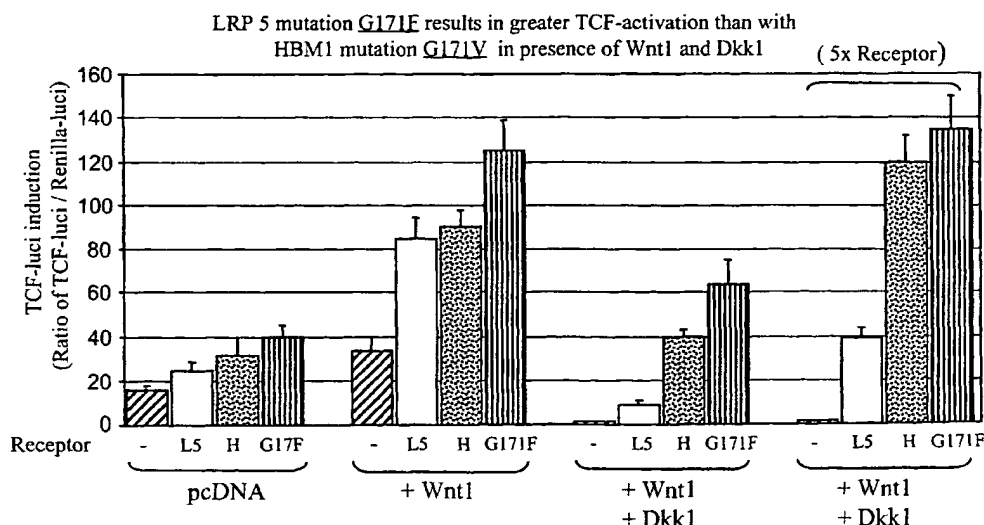
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ALLEN, Kristina**

[Continued on next page]

(54) Title: HBM VARIANTS THAT MODULATE BONE MASS AND LIPID LEVELS



* G171F mutation involves the ringed R group (F) alteration and leads to marginally greater TCF-luciferase activation than that with HBM1 mutation G171V.

(57) Abstract: ABSTRACT The present invention relates to methods and materials used to express an HBM-like polypeptide derived from HBM, LRP5 or LRP6 in animal cells and transgenic animals. The present invention also relates to transgenic animals expressing the HBM-like polypeptides. The invention provides nucleic acids, including coding sequences, oligonucleotide primers and probes, proteins, cloning vectors, expression vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in bone development, and methods of diagnosing and treating diseases involved in bone development and lipid modulation. In preferred embodiments, the present invention is directed to methods for treating, diagnosing and preventing osteoporosis.



(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— without international search report and to be republished upon receipt of that report

HBM VARIANTS THAT MODULATE BONE MASS AND LIPID LEVELS

FIELD OF THE INVENTION

The present invention relates generally to the field of genetics, genomics and molecular biology. The invention relates to methods and materials used to isolate, detect and sequence a high bone mass gene and corresponding wild-type gene, and mutants thereof. The present invention also relates to the high bone mass (HBM) gene, the corresponding wild-type gene, and mutants thereof. The genes identified in the present invention are implicated in the ontology and physiology of bone development. The invention also provides nucleic acids, proteins, cloning vectors, expression vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in bone development, and methods of diagnosing and treating diseases involved in bone development and lipid levels. The invention further relates to transgenic animals for studying the HBM phenotype and related variant phenotypes, the mechanism of action of the *HBM* gene and its variants, and factors and treatments affecting normal and abnormal bone conditions. In preferred embodiments, the present invention is directed to methods for treating, diagnosing, preventing and screening for normal and abnormal conditions of bone, including metabolic bone diseases such as osteoporosis.

BACKGROUND OF THE INVENTION

Two of the most common types of osteoporosis are postmenopausal and senile osteoporosis. Osteoporosis affects men as well as women, and, taken with other abnormalities of bone, presents an ever-increasing health risk for an aging population. The most common type of osteoporosis is that associated with menopause. Most women lose between 20-60% of the bone mass in the trabecular compartment of the bone within 3-6 years after the cessation of menses. This rapid loss is generally associated with an increase of bone resorption and formation. However, the resorptive cycle is more dominant and the result is a net loss of bone mass. Osteoporosis is a common and serious disease among postmenopausal women. There are an estimated 25 million women in the United States alone who are afflicted with this disease. The results of osteoporosis are both personally harmful, and also

account for a large economic loss due to its chronicity and the need for extensive and long-term support (hospitalization and nursing home care) from the disease sequelae. This is especially true in more elderly patients. Additionally, while osteoporosis is generally not thought of as a life-threatening condition, a 20-30% mortality rate is related to hip fractures in elderly women. A large percentage of this mortality rate can be directly associated with postmenopausal osteoporosis. The costs alone associated with the treatment of osteoporotic fractures in the United States is \$10 to \$15 billion annually. Worldwide incidence of osteoporotic hip fractures is estimated to exceed 1.7 million cases.

The most vulnerable tissue in the bone to the effects of postmenopausal osteoporosis is the trabecular bone. This tissue is often referred to as spongy bone and is particularly concentrated near the ends of the bone near the joints and in the vertebrae of the spine. The trabecular tissue is characterized by small structures which inter-connect with each other as well as the more solid and dense cortical tissue which makes up the outer surface and central shaft of the bone. This crisscross network of trabeculae gives lateral support to the outer cortical structure and is critical to the biomechanical strength of the overall structure. In postmenopausal osteoporosis, it is primarily the net resorption and loss of the trabeculae which lead to the failure and fracture of the bone. In light of the loss of the trabeculae in postmenopausal women, it is not surprising that the most common fractures are those associated with bones which are highly dependent on trabecular support, e.g., the vertebrae, the neck of the femur, and the forearm. Indeed, hip fracture, Colle's fractures, and vertebral crush fractures are indicative of postmenopausal osteoporosis.

One of the earliest generally accepted methods for treatment of postmenopausal osteoporosis was estrogen replacement therapy. Although this therapy frequently is successful, patient compliance is low, primarily due to the undesirable side-effects of chronic estrogen treatment. Frequently cited side-effects of estrogen replacement therapy include reinitiation of menses, bloating, depression, and fear of breast or uterine cancer. In order to limit the known threat of uterine cancer in those women who have not undergone a hysterectomy, a protocol of estrogen and progestin cyclic therapy is often employed. This protocol is similar to that which is used in birth control regimens, and often is not tolerated by women because of the side-effects characteristic of progestin. More recently, certain

antiestrogens, originally developed for the treatment of breast cancer, have been shown in experimental models of postmenopausal osteoporosis to be efficacious. Among these agents is raloxifene (See, U.S. Patent No. 5,393,763, and Black et al, *J. Clin. Invest.*, 93:63-69 (1994)). In addition, tamoxifene, a widely used clinical agent for the treatment of breast cancer, has been shown to increase bone mineral density in post menopausal women suffering from breast cancer (Love et al, *N. Engl. J. Med.*, 326:852-856 (1992)).

Another therapy for the treatment of postmenopausal osteoporosis is the use of calcitonin. Calcitonin is a naturally occurring peptide which inhibits bone resorption and has been approved for this use in many countries (Overgaard et al, *Br. Med. J.*, 305:556-561 (1992)). The use of calcitonin has been somewhat limited, however. Its effects are very modest in increasing bone mineral density and the treatment is very expensive. Another therapy for the treatment of postmenopausal osteoporosis is the use of bis-phosphonates. These compounds were originally developed for use in Paget's disease and malignant hypercalcemia. They have been shown to inhibit bone resorption. Alendronate, one compound of this class, has been approved for the treatment of postmenopausal osteoporosis. These agents may be helpful in the treatment of osteoporosis, but these agents also have potential liabilities which include osteomalacia, extremely long half-life in bone (greater than 2 years), and possible "frozen bone syndrome," e.g., the cessation of normal bone remodeling.

Senile osteoporosis is similar to postmenopausal osteoporosis in that it is marked by the loss of bone mineral density and resulting increase in fracture rate, morbidity, and associated mortality. Generally, it occurs in later life, i.e., after 70 years of age. Historically, senile osteoporosis has been more common in females, but with the advent of a more elderly male population, this disease is becoming a major factor in the health of both sexes. It is not clear what, if any, role hormones such as testosterone or estrogen have in this disease, and its etiology remains obscure. Treatment of this disease has not been very satisfactory. Hormone therapy, estrogen in women and testosterone in men, has shown equivocal results; calcitonin and bis-phosphonates may be of some utility.

The peak mass of the skeleton at maturity is largely under genetic control. Twin studies have shown that the variance in bone mass between adult monozygotic twins is

smaller than between dizygotic twins (Slemenda et al, *J. Bone Miner. Res.*, 6:561-567 (1991); Young et al, *J. Bone Miner. Res.*, 6:561-567 (1995); Pocock et al, *J. Clin. Invest.*, 80:706-710 (1987); Kelly et al, *J. Bone Miner. Res.*, 8:11-17 (1993)), and it has been estimated that up to 60% or more of the variance in skeletal mass is inherited (Krall et al, *J. Bone Miner. Res.*, 10:S367 (1993)). Peak skeletal mass is the most powerful determinant of bone mass in elderly years (Hui et al, *Ann. Int. Med.*, 111:355-361 (1989)), even though the rate of age-related bone loss in adult and later life is also a strong determinant (Hui et al, *Osteoporosis Int.*, 1:30-34 (1995)). Since bone mass is the principal measurable determinant of fracture risk, the inherited peak skeletal mass achieved at maturity is an important determinant of an individual's risk of fracture later in life. Thus, study of the genetic basis of bone mass is of considerable interest in the etiology of fractures due to osteoporosis.

Recently, a strong interest in the genetic control of peak bone mass has developed in the field of osteoporosis. The interest has focused mainly on candidate genes with suitable polymorphisms to test for association with variation in bone mass within the normal range, or has focused on examination of genes and gene loci associated with low bone mass in the range found in patients with osteoporosis. The vitamin D receptor locus (VDR) (Morrison et al, *Nature*, 367:284-287 (1994)), PTH gene (Howard et al, *J. Clin. Endocrinol. Metab.*, 80:2800-2805 (1995); Johnson et al, *J. Bone Miner. Res.*, 8:11-17 (1995); Gong et al, *J. Bone Miner. Res.*, 10:S462 (1995)) and the estrogen receptor gene (Hosoi et al, *J. Bone Miner. Res.*, 10:S170 (1995); Morrison et al, *Nature*, 367:284-287 (1994)) have figured most prominently in this work. These studies are difficult because bone mass (the phenotype) is a continuous, quantitative, polygenic trait, and is confounded by environmental factors such as nutrition, co-morbid disease, age, physical activity, and other factors. Also, this type of study design requires large numbers of subjects. In particular, the results of VDR studies to date have been confusing and contradictory (Garnero et al, *J. Bone Miner. Res.*, 10:1283-1288 (1995); Eisman et al, *J. Bone Miner. Res.*, 10:1289-1293 (1995); Peacock, *J. Bone Miner. Res.*, 10:1294-1297 (1995)). Furthermore, the work thus far has not shed much light on the mechanism(s) whereby the genetic influences might exert their effect on bone mass.

While it is well known that peak bone mass is largely determined by genetic rather than environmental factors, studies to determine the gene loci (and ultimately the genes)

linked to variation in bone mass are difficult and expensive. Study designs which utilize the power of linkage analysis, e.g., sib-pair or extended family, are generally more informative than simple association studies, although the latter do have value. However, genetic linkage studies involving bone mass are hampered by two major problems. The first problem is the phenotype, as discussed briefly above. Bone mass is a continuous, quantitative trait, and establishing a discrete phenotype is difficult. Each anatomical site for measurement may be influenced by several genes, many of which may be different from site to site. The second problem is the age component of the phenotype. By the time an individual can be identified as having low bone mass, there is a high probability that their parents or other members of prior generations will be deceased and therefore unavailable for study, and younger generations may not have even reached peak bone mass, making their phenotyping uncertain for genetic analysis.

Regardless, linkage analysis can be used to find the location of a gene causing a hereditary "disorder" and does not require any knowledge of the biochemical nature of the disorder, i.e., a mutated protein that is believed to cause the disorder does not need to be known. Traditional approaches depend on assumptions concerning the disease process that might implicate a known protein as a candidate to be evaluated. The genetic localization approach using linkage analysis can be used to first find the general chromosomal region in which the defective gene is located and then to gradually reduce the size of the region in order to determine the location of the specific mutated gene as precisely as possible. After the gene itself is discovered within the candidate region, the messenger RNA and the protein are identified and, along with the DNA, are checked for mutations.

The genetic localization approach has practical implications since the location of the disease can be used for prenatal diagnosis even before the altered gene that causes the disease is found. Linkage analysis can enable families, even many of those that do not have a sick child, to know whether they are carriers of a disease gene and to evaluate the condition of an unborn child through molecular diagnosis. The transmission of a disease within families, then, can be used to find the defective gene. As used herein, reference to "high bone mass" (HBM) is analogous to reference to a disease state, although from a practical standpoint high bone mass can actually help a subject avoid the disease known as osteoporosis.

Linkage analysis is possible because of the nature of inheritance of chromosomes from parents to offspring. During meiosis, the two parental homologues pair to guide their proper separation to daughter cells. While they are lined up and paired, the two homologues exchange pieces of the chromosomes, in an event called "crossing over" or "recombination."

5 The resulting chromosomes are chimeric, that is, they contain parts that originate from both parental homologues. The closer together two sequences are on the chromosome, the less likely that a recombination event will occur between them, and the more closely linked they are. In a linkage analysis experiment, two positions on the chromosomes are followed from one generation to the next to determine the frequency of recombination between them. In a
10 study of an inherited disease, one of the chromosomal positions is marked by the disease gene or its normal counterpart, i.e., the inheritance of the chromosomal region can be determined by examining whether the individual displays symptoms of the disorder or not. The other position is marked by a DNA sequence that shows natural variation in the population such that the two homologues can be distinguished based on the copy of the "marker" sequence
15 that they possess. In every family, the inheritance of the genetic marker sequence is compared to the inheritance of the disease state. If, within a family carrying an autosomal dominant disorder such as high bone mass, every affected individual carries the same form of the marker and all the unaffected individuals carry at least one different form of the marker, there is a great probability that the disease gene and the marker are located close to each
20 other. In this way, chromosomes may be systematically checked with known markers and compared to the disease state. The data obtained from the different families is combined, and analyzed together by a computer using statistical methods. The result is information indicating the probability of linkage between the genetic marker and the disease allowing different distances between them. A positive result can mean that the disease is very close to
25 the marker, while a negative result indicates that it is far away on that chromosome, or on an entirely different chromosome.

Linkage analysis is performed by typing all members of the affected family at a given marker locus and evaluating the co-inheritance of a particular disease state with the marker probe, thereby determining how often the two of them are co-inherited. The recombination
30 frequency can be used as a measure of the genetic distance between two gene loci. A

recombination frequency of 1% is equivalent to 1 map unit, or 1 centiMorgan (cM), which is roughly equivalent to 1,000 kb of DNA. This relationship holds up to frequencies of about 20% or 20 cM.

The entire human genome is 3,300 cM long. In order to find an unknown disease gene within 5-10 cM of a marker locus, the whole human genome can be searched with roughly 330 informative marker loci spaced at approximately 10 cM intervals (Botstein *et al.*, *Am. J. Hum. Genet.*, 32:314-331 (1980)). The reliability of linkage results is established by using a number of statistical methods. The method most commonly used for the analysis of linkage in humans is the LOD score method (Morton, *Prog. Clin. Biol. Res.*, 147:245-265 (1984), Morton *et al.*, *Am. J. Hum. Genet.*, 38:868-883 (1986)) which was incorporated into the computer program, LIPED, by Ott, *Am. J. Hum. Genet.*, 28:528-529 (1976). LOD scores are the logarithm of the ratio of the likelihood that two loci are linked at a given distance to that they are not linked (>50 cM apart). The advantage of using logarithmic values is that they can be summed among families with the same disease. This becomes necessary given the relatively small size of human families.

By convention, a total LOD score greater than + 3.0 (that is, odds of linkage at the specified recombination frequency being 1000 times greater than odds of no linkage) is considered to be significant evidence for linkage at that particular recombination frequency. A total LOD score of less than - 2.0 (that is, odds of no linkage being 100 times greater than odds of linkage at the specified frequency) is considered to be strong evidence that the two loci under consideration are not linked at that particular recombination frequency. Until recently, most linkage analyses have been performed on the basis of two-point data, which is the relationship between the disorder under consideration and a particular genetic marker. However, as a result of the rapid advances in mapping the human genome over the last few years, and concomitant improvements in computer methodology, it has become feasible to carry out linkage analyses using multi-point data. Multi-point analysis provide a simultaneous analysis of linkage between the disease and several linked genetic markers, when the recombination distance among the markers is known.

Multi-point analysis is advantageous for two reasons. First, the informativeness of the pedigree is usually increased. Each pedigree has a certain amount of potential information,

dependent on the number of parents heterozygous for the marker loci and the number of affected individuals in the family. However, few markers are sufficiently polymorphic as to be informative in all those individuals. If multiple markers are considered simultaneously, then the probability of an individual being heterozygous for at least one of the markers is greatly increased. Second, an indication of the position of the disease gene among the markers may be determined. This allows identification of flanking markers, and thus eventually allows isolation of a small region in which the disease gene resides. Lathrop *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:3443-3446 (1984) have written the most widely used computer package, *LINKAGE*, for multi-point analysis.

There is a need in the art for identifying the gene associated with a high bone mass phenotype. There is also a need for tools for the study of the high bone mass gene and phenotype. More generally there is need for the development of diagnostic tools and treatments. The present invention is directed to these, as well as other, important ends.

SUMMARY OF THE INVENTION

It is an object of the invention to provide a nucleic acid comprising a mutation in *LRP5* or *LRP6* which results in a HBM phenotype when expressed in a cell, wherein said HBM phenotype results in bone mass modulation and/or lipid level modulation. Another embodiment contemplates that the mutation is located in propeller 1. In another embodiment, the nucleic acid encodes a mutation comprising at least one mutation of Tables 2 or 3 or which results in a mutation of one of the following G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V when expressed in a cell, and wherein expression of the nucleic acid in a subject results in bone mass modulation and/or lipid level modulation. In the instance of *LRP6*, the mutation is located in a position equivalent to *LRP5* such that expression of the nucleic acid in a subject results in bone mass modulation and/or lipid level modulation. In another embodiment, the preferred mutation of *LRP5* is G171V, A214V, A65V, M282V, G171K, G171F, G171I or G171Q or in an equivalent location if dealing with *LRP6*.

Another embodiment contemplated herein is a polypeptide encoded by any of the above nucleic acids, wherein said polypeptide when expressed in a cell modulates Wnt

signaling, LRP5 activity and/or LRP6 activity. The polypeptide can additionally or alternatively modulate bone mass and/or lipid levels when expressed in a subject. These polypeptides or biologically active fragments thereof may preferably contain any of the following mutations of Table 2, G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V in LRP5 or in a equivalent location in LRP6. The most preferred mutations are G171V, A214V, A65V, M282V, G171K, G171F, G171I or G171Q in LRP5 or an equivalent location in LRP6.

Yet another embodiment contemplates antibodies an immunogenic fragments thereof which bind to these proteins. The contemplated antibodies include a monoclonal antibody, a chimeric antibody, a bispecific antibody, a humanized antibody, a primatized® antibody, a human antibody, or a labeled antibody. Preferably, some of the antibodies can discriminate between the wild type and variant forms of LRP5 and LRP6.

Another embodiment contemplates antibodies which bind to polypeptides comprising:
²⁰⁸KLYWADAKLSFIHRAN²²³, ²⁷⁷ALYSPMDIQVLSQER²⁹¹,
⁶¹GLEDAAAVDFQFSKGA⁷³, ²³⁴EGSLTHPFALTLTG²⁴⁷, ²⁴⁹TLYWTDWQTRSIHACN²⁶⁴,
¹⁴⁴VLFWQDLDPRAI¹⁵⁶, ¹⁹⁴IYWPNGLTIDLEEQKLY²¹⁰, ³⁴LLLFANRRDVRLVD⁴⁷,
⁷⁵GAVYWTDVSEEAIKQ⁸⁹, ¹²¹KLYWTDSETNRIEVA¹³⁵ of LRP5 or an equivalent domain on LRP6 or variants thereof.

The above antibodies can also be used in a composition for modulating bone mass and/or lipid levels in a subject comprising a therapeutically effective amount of the antibody or immunogenic fragment and a pharmaceutically acceptable carrier.

The invention further contemplates a method of diagnosing a HBM like phenotype in a subject comprising: (A) obtaining a biological sample from said subject; (B) exposing the sample to one of the described antibodies or immunogenic fragments; and (C) detecting whether the antibody bound a protein from the biological sample from said subject to determine whether the subject has a HBM-like phenotype.

Another embodiment contemplates a transgenic animal having somatic and/or germ cells comprising a nucleic acid which comprises a promoter region that directs protein expression in animal and/or human cells operably linked to any of the herein described nucleic acids, and wherein said transgenic animal has at least three bone parameters

modulated by the expression of said nucleic acid. The promoter region can be selected from the group consisting of CMV, RSV, SV40, and EF-1a, CMV β Actin, histone, type I collagen, TGF β 1, SX2, cfos/cjun, Cbfa1, Fra/Jun, Dlx5, osteocalcin, osteopontin, bone sialoprotein, and collagenase promoter regions.

5 A further embodiment of the invention contemplates an animal model for the study of bone density modulation and/or lipid level modulation comprising a first group of animals composed of any of the described transgenic animals and a second group of control animals.

Another embodiment provides for a method of identifying agents which modulate the activity of an HBM-like nucleic acid comprising: (A) transfecting a cell with a vector of
10 claim 11; (B) exposing the transfected cell of step (A) to a compound; and (C) determining whether the compound modulates the activity of the HBM-like nucleic acid. Such agents can include a hormone, a growth factor, a peptide, RNA, siRNA, DNA, a mineral, a vitamin, a natural product, or a synthetic organic compound.

Another aspect of the invention provides for a method for identifying compounds
15 which modulate the interaction of Dkk with the Wnt signaling pathway comprising: (A) transfecting cells with constructs containing any of the described nucleic acids; (B) assessing changes in expression of a reporter element linked to a Wnt-responsive promoter; and (C) identifying as a Dkk/Wnt interaction modulating compound any compound which alters reporter gene expression compared with cells transfected with a Dkk construct alone. The
20 cells are preferably cancer cells, liver cells or bone cells. The reporter element used is TCF-luciferase, tk-Renilla, or a combination thereof.

Yet another embodiment includes a method of diagnosing a subject as expressing a nucleic acid comprising a nucleotide change of Tables 2 or 3 or any other mutations, the method comprising the steps of: (A) obtaining a biological sample from the subject; and (B)
25 assaying for the presence of the nucleotide change which results in HBM phenotype.

The invention also provides agents identified by the above methods which regulate Wnt activity, Dkk activity, bone mass and/or lipid levels.

BRIEF DESCRIPTION OF THE FIGURES

30 **Fig. 1** shows the pedigree of the individuals used in the genetic linkage studies.

Under each individual is an ID number, the z-score for spinal BMD, and the allele calls for the critical markers on chromosome 11. Solid symbols represent "affected" individuals. Symbols containing "N" are "unaffected" individuals. DNA from 37 individuals was genotyped. Question marks denote unknown genotypes or individuals who were not
 5 genotyped.

Fig. 2 depicts the BAC/STS content physical map of the HBM region in 11q13.3. STS markers derived from genes, ESTs, microsatellites, random sequences, and BAC end sequences are denoted above the long horizontal line. For markers that are present in GDB the same nomenclature has been used. Locus names (D11S####) are listed in parentheses
 10 after the primary name if available. STSs derived from BAC end sequences are listed with the BAC name first followed by L or R for the left and right end of the clone, respectively. The two large arrows indicate the genetic markers that define the HBM critical region. The horizontal lines below the STSs indicate BAC clones identified by PCR-based screening of a nine-fold coverage BAC library. Open circles indicate that the marker did not amplify the
 15 corresponding BAC library address during library screening. Clone names use the following convention: B for BAC, the plate, row and column address, followed by -H indicating the HBM project (i.e., B36F16-H).

Figs. 3A-3F show the genomic structure of Zmax1 (LRP5) with flanking intron sequences. Translation is initiated by the underlined "ATG" in exon 1. The site of the
 20 polymorphism in the *HBM* gene is in exon 3 and is represented by the underlined "G," whereby this nucleotide is a "T" in the *HBM* gene. The 3' untranslated region of the mRNA is underlined within exon 23 (exon 1, SEQ ID NO:40; exon 2, SEQ ID NO:41; exon 3, SEQ ID NO:42; exon 4, SEQ ID NO:43; exon 5, SEQ ID NO:44; exon 6, SEQ ID NO:45; exon 7, SEQ ID NO:46; exon 8, SEQ ID NO:47; exon 9, SEQ ID NO:48; exon 10, SEQ ID NO:49;
 25 exon 11, SEQ ID NO:50; exon 12, SEQ ID NO:51; exon 13, SEQ ID NO:52; exon 14, SEQ ID NO:53; exon 15, SEQ ID NO:54; exon 16, SEQ ID NO:55; exon 17, SEQ ID NO:56; exon 18, SEQ ID NO:57; exon 19, SEQ ID NO:58; exon 20, SEQ ID NO:59; exon 21, SEQ ID NO:60; exon 22, SEQ ID NO:61; and exon 23; SEQ ID NO:62).

Fig. 4 shows the domain organization of Zmax1 (LRP5), including the YWTD
 30 spacers, the extracellular attachment site, the binding site for LDL and calcium, the cysteine-

rich growth factor repeats, the transmembrane region, the ideal PEST region with the CK-II phosphorylation site and the internalization domain. Fig. 4 also shows the site of the glycine to valine change that occurs in the HBM protein. The signal peptide is located at amino acids 1-31, the extracellular domain is located at amino acids 32-1385, the transmembrane segment is located at amino acids 1386-1413, and the cytoplasmic domain is located at amino acids 1414-1615.

Fig. 5 is a schematic illustration of the BAC contigs B527D12 and B200E21 in relation to the *HBM* gene.

Figs. 6A-6J are the nucleotide (SEQ ID NO:1) and amino acid (SEQ ID NO:3) sequences of the wild-type gene, *Zmax1* (*LRP5*). The location for the base pair substitution at nucleotide 582, a guanine to thymine, (SEQ ID NO: 2 and 4) is underlined. This allelic variant is the *HBM* gene. The *HBM* gene encodes for a protein with an amino acid substitution of glycine to valine at position 171. The 5' untranslated region (UTR) boundaries bases 1 to 70, and the 3' UTR boundaries bases 4916-5120.

Figs. 7A and 7B are northern blot analyses showing the expression of *Zmax1* (*LRP5*) in various tissues.

Fig. 8 is a PCR product analysis.

Fig. 9 is allele specific oligonucleotide detection of the *Zmax1* (*LRP5*) exon 3 mutation.

Fig. 10 is the cellular localization of mouse *Zmax1* (*LRP5*) by *in situ* hybridization at 100X magnification using sense and antisense probes.

Fig. 11 is the cellular localization of mouse *Zmax1* (*LRP5*) by *in situ* hybridization at 400X magnification using sense and antisense probes.

Fig. 12 is the cellular localization of mouse *Zmax1* (*LRP5*) by *in situ* hybridization of osteoblasts in the endosteum at 400X magnification using sense and antisense probes.

Fig. 13 shows antisense inhibition of *Zmax1* (*LRP5*) expression in MC-3T3 cells.

Fig. 14 shows a *Zmax1* (*LRP5*) exon 3 allele specific oligonucleotide (ASO) assay which illustrates the rarity of the HBM allele (right panels; T-specific oligo; 58°C Wash) as compared to the wild-type *Zmax1* (*LRP5*) allele (left panels, G-specific oligo; 55°C Wash). The positive spots appearing in the right panels were positive controls.

Fig. 15 depicts a model representing the potential role of Zmax1 (LRP5) in focal adhesion signaling.

Fig. 16 depicts a schematic of two *Zmax1* (LRP5) gene targeting vectors for the knock-out of endogenous mouse Zmax or conditional knock-in of the HBM polymorphism. B, X, and R indicate *Bam*HI, *Xba*I, and *Eco*RI sites in DNA BAC 4735P5 respectively. Exons 3, 4, and 5 are indicated by black rectangles. A G → T base change is engineered at base 24 of exon 3 to produce the HBM polymorphism. The location of a LoxP flanked cassette containing a neomycin resistance gene and a synthetic pause sequence and probes used for screening and characterizing of ES cell clones are also indicated.

Fig. 17 confirms expression by the transgenic (*i.e.*, HBMMCBA and HBMMTIC) and wild-type (*i.e.*, ZmaxWTCBA and ZmaxWTTIC) plasmid constructs. These constructs were transiently transfected into HOB-02-02 cells and the mRNA levels determined using TaqMan® quantitative PCR. HBMMCBA and ZmaxWTCBA are shown in the left column (*i.e.* CMVβActin) and HBMMTIC and ZmaxWTTIC are shown in the right column (*i.e.* Type I collagen) of the Table.

Fig. 18 depicts a comparison between the human and mouse TaqMan® Primer/Probe sets. HOB (HOB-03-C5) and mouse (MC-3T3-E1) osteoblastic cell mRNA was analyzed using the probes and primers.

Fig. 19 depicts the quantification of human Zmax1 (LRP5) mRNA expressed in a mixed human and mouse RNA background using the TaqMan® Primer/Probe sets. Results are presented in Human Zmax1 (LRP5) mRNA added (ng) versus Human Zmax1 (LRP5) mRNA measured (ng).

Fig. 20 depicts expression of HBM in transgenic mice based on mRNA expression analyzed by TaqMan®.

Fig. 21A-F Panels A-C depict the analysis of various transgenic mouse lines which express the HBMMCBA construct in spine (Fig. 21A), femur (Fig. 21B) and total body (Fig. 21C). Panels D-F depict the analysis of various transgenic mouse lines which express the HBMMTIC construct in spine (Fig. 21D), femur (Fig. 21E) and total body (Fig. 21F).

Fig. 22 depicts changes in BMD in HBM transgenic mice (*i.e.*, HBMMCBA and HBMMTIC constructs) at 5 weeks using *in vivo* pDXA* analysis. The BMD changes are

presented as compared to wild-type animals which were also only 5 weeks old.

Fig. 23 depicts changes in BMD in HBM transgenic mice (*i.e.*, HBMMCBA and HBMMTIC constructs) at 9 weeks using *in vivo* pDXA* analysis. The BMD changes are presented as compared to wild-type animals which were also only 9 weeks old.

5 **Fig. 24** (A-D) presents the sequence of the insert of the gene, HBMGI_2AS, subcloned into the vector (SEQ ID NO:759).

Fig. 25 (A-D) presents the sequence of the insert of the gene, ZMAXGI_3AS, subcloned into the vector (SEQ ID NO:760).

10 **Fig. 26** (A-C) presents an alignment of human (SEQ ID NO:761) and mouse (SEQ ID NO:762) Zmax1 (LRP5) amino acid sequences.

Fig. 27 (A-C) presents an alignment of human LRP5 (Zmax1) (SEQ ID NO:763) and LRP6 (SEQ ID NO:764) amino acid sequences.

Fig. 28 shows a schematic of the components of the Wnt signal transduction pathway. Schematic obtained from: <http://www.stanford.edu/~rnusse/pathways/cell2.html>.

15 **Fig. 29** shows that the mutation G171F in LRP5 produces a greater activation of the Wnt pathway than LRP5 which is consistent with HBM activity.

Fig. 30 shows that the mutation M282V in LRP5 produces an activation of the Wnt pathway which is consistent with HBM activity in U2OS cells.

20 **Fig. 31** shows a table of proteins identified in a Y2H screen using a Dkk-1 bait sequence. These proteins are identified by both their nucleic acid and amino acid accession numbers.

Fig. 32 shows the differential binding of an antibody generated to a sequence (a.a. 165-177) containing the HBM mutation in LRP5 in LRP5 and HBM virus-infected cells.

Fig. 33 shows a diagram of the *Xenopus* Embryo Assay for Wnt activity.

25 **Fig. 34** shows the effects of Zmax/LRP5 and HBM on Wnt signaling in the *Xenopus* embryo assay.

Fig. 35 shows the effects of Zmax/LRP5 and HBM on induction of secondary axis formation in the *Xenopus* embryo assay.

30 **Fig. 36** shows the effects of human Dkk-1 on the repression of the canonical Wnt pathway.

Fig. 37 shows the effects of human Dkk-1 on Zmax/LRP5 and HBM-mediated Wnt signaling.

Fig. 38 shows a table of peptide aptamer insert sequences identified in a Y2H screen using a LRP5 ligand binding domain bait sequence.

5 **Fig. 39** shows pcDNA3.1 construct names with nucleotide sequences for LRP5-binding peptide aptamers, Dkk-1 peptides and control constructs.

10 **Fig. 40** shows the results of a minimum interaction domain mapping screen of Dkk-1 with LRP5. At the top, a map of Dkk-1 showing the location of the signal sequence, and cysteine rich domains 1 and 2. Below, the extent of domains examined using LRP5 LBD baits, LBD1 and LBD4. To the right, scoring of the binding results observed in the experiment.

Fig. 41 shows the effects of Dkk-1 and Dkk-2 on Wnt1 signaling with coreceptors LRP5, HBM, and LRP6 in HOB03CE6 cells.

15 **Fig. 42** shows the effects of Dkk-1 and Dkk-2 on Wnt3a signaling with coreceptors LRP5, HBM, and LRP6 in HOB03CE6 cells.

Fig. 43 demonstrates that the LRP5-LBD peptide aptamer 262 activates Wnt signaling in the presence of Wnt3a in U2OS cells.

Fig. 44 shows the amino acid sequences for the corresponding LRP5-binding peptides, Dkk-1 peptide aptamers and control constructs in Figure 39.

20 **Fig. 45** shows that Dkk-1 represses Wnt3a-mediated Wnt signaling in U2OS bone cells using the cell-based reporter gene assay for high throughput screening.

Fig. 46 demonstrates that Wnt1-HBM generated signaling is not efficiently inhibited by Dkk-1 in U2OS bone cells while LRP5 and LRP6-mediated signaling are using the cell-based reporter gene assay for high throughput screening.

25 **Fig. 47** shows that the TCF signal in the cell-based reporter gene assay for high throughput screening can be modulated by Dkk-1 and Dkk-1-AP without Wnt DNA transfection.

Fig. 48 shows the morphological results in the Xenopus assay using aptamers 261 and 262 from the LRP5-LBD to activate Wnt signaling.

30 **Fig. 49** demonstrates that LRP5-LBD aptamers 261 and 262 induce Wnt signaling

over other LRP5 aptamers.

Fig. 50 depicts the LRP5 domain structure. The symbols are defined at the upper right. The structural domains are numbered successively from top to bottom corresponding to the N-terminal to C-terminal ends of the protein as Propeller 1, EGF-like Domain 1, Propeller 2, EGF-like Domain 2, and so forth. The structure was determined using the motifs and nomenclature described in Springer *et al.*, 1998 *J. Mol. Biol.* 283: 837-62.

Fig. 51 CLUSTALW (1.8) multiple sequence alignment of the β -propeller segments of LRP5 from human (af077820h) and mouse (af064984m and af077847m), LRP6 from human (af074264h) and mouse (af074265m), and the sequences from proteins with modeled 6-bladed β -propellers: chicken LRP1(1lpx) and human nidogen (1ndx). For the LPRs, the final suffix letter designates which of the four propeller domains the sequence comes from (a => prop. #1, b => prop. #2, etc.). The final four lines give the secondary structure assignments predicted by DSC for all sequences given in this alignment (H = helix, E = strand and C = coil) and the weights assigned to each structural type at each position on a scale of 0 (least probable) to 9 (most probable). Sequences corresponding to Springer's models' strand positions are underlined and color-coded according to which of the six blades they belong to. In the propeller geometry, alternate loops fall on opposite faces of the disc-shaped domain. The loops on the "top" face (*i.e.*, opposite the points of entry and exit of the chain from the structure) are colored red. The position of the G171V mutation is marked with "*".

Fig. 52 displays the functional effect of mutations on side chain interactions in HBM protein as compared the wild-type Zmax1 (LRP5) protein. It also shows the side chain interactions of G171F, an HBM like variant.

Fig. 53 displays the structural change due to the OPPG mutation versus the wild-type, Zmax1 (LRP5). Panel A depicts the homology model of LRP5's second propeller domain; Panel B depicts the third propeller domain and Panel C depicts the second propeller domain with two mutations.

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

To aid in the understanding of the specification and claims, the following definitions are provided.

5 "Gene" refers to a DNA sequence that encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide. The term "gene" includes intervening, non-coding regions, as well as regulatory regions, and can include 5' and 3' ends.

10 By "nucleic acid" is meant to include single stranded and double stranded nucleic acids including, but not limited to DNAs, RNAs (e.g., mRNA, tRNAs, siRNAs), cDNAs, recombinant DNA (rDNA), rRNAs, antisense nucleic acids, oligonucleotides, and oligomers, and polynucleotides. The term may also include hybrids such as triple stranded regions of RNA and/or DNA or double stranded RNA:DNA hybrids. The term also is contemplated to include modified nucleic acids such as, but not limited to biotinylated nucleic acids, tritylated
15 nucleic acids, fluorophor labeled nucleic acids, inosine, and the like.

"Gene sequence" refers to a nucleic acid molecule, including DNA which contains a non-transcribed or non-translated sequence, which comprises a gene. The term is also intended to include any combination of gene(s), gene fragment(s), non-transcribed sequence(s) or non-translated sequence(s) which are present on the same DNA molecule.

20 The nucleic acid sequences of the present invention may be derived from a variety of sources including DNA, cDNA, synthetic DNA, synthetic RNA or combinations thereof. Such sequences may comprise genomic DNA which may or may not include naturally occurring introns. Moreover, such genomic DNA may be obtained in association with promoter regions and/or poly (A) sequences. The sequences, genomic DNA or cDNA may
25 be obtained in any of several ways. Genomic DNA can be extracted and purified from suitable cells by means well known in the art. Alternatively, mRNA can be isolated from a cell and used to produce cDNA by reverse transcription or other means.

"cDNA" refers to complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus, a "cDNA
30 clone" means a duplex DNA sequence for which one strand is complementary to an RNA

molecule of interest, carried in a cloning vector or PCR amplified. cDNA can also be single stranded after first strand synthesis by reverse transcriptase. In this form, it is a useful PCR template and does not need to be carried in a cloning vector. This term includes genes from which the intervening sequences have been removed. Thus, the term "gene", as sometimes
5 used generically, can also include nucleic acid molecules comprising cDNA and cDNA clones.

"Recombinant DNA" means a molecule that has been engineered by splicing *in vitro* a cDNA or genomic DNA sequence or altering a sequence by methods such as PCR mutagenesis.

10 "Cloning" refers to the use of *in vitro* recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired gene, it is necessary to use methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient
15 host cells.

"cDNA library" refers to a collection of recombinant DNA molecules containing cDNA inserts which together comprise the entire or a partial repertoire of genes expressed in a particular tissue or cell source. Such a cDNA library can be prepared by methods known to one skilled in the art and described by, for example, Cowell and Austin, "cDNA Library
20 Protocols," *Methods in Molecular Biology* (1997).

"Cloning vehicle" refers to a plasmid or phage DNA or other DNA sequence which is able to replicate in a host cell. This term can also include artificial chromosomes such as BACs and YACs. The cloning vehicle is characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without
25 loss of an essential biological function of the DNA, which may contain a marker suitable for use in the identification of transformed cells.

"Expression" refers to the process comprising transcription of a gene sequence and subsequent processing steps, such as translation of a resultant mRNA to produce the final end product of a gene. The end product may be a protein (such as an enzyme or receptor) or a
30 nucleic acid (such as a tRNA, antisense RNA, or other regulatory factor). The term

"expression control sequence" refers to a sequence of nucleotides that control or regulate expression of structural genes when operably linked to those genes. These include, for example, the lac systems, the trp system, major operator and promoter regions of the phage lambda, the control region of fd coat protein and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells. Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host, and may contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements and/or translational initiation and termination sites.

"Expression vehicle" refers to a vehicle or vector similar to a cloning vehicle but which is capable of expressing a gene which has been cloned into it, after transformation into a host. The cloned gene is usually placed under the control of (i.e., operably linked to) an expression control sequence.

"Operator" refers to a DNA sequence capable of interacting with the specific repressor, thereby controlling the transcription of adjacent gene(s).

"Promoter" refers to a DNA sequence that can be recognized by an RNA polymerase. The presence of such a sequence permits the RNA polymerase to bind and initiate transcription of operably linked gene sequences.

"Promoter region" is intended to include the promoter as well as other gene sequences which may be necessary for the initiation of transcription. The presence of a promoter region is sufficient to cause the expression of an operably linked gene sequence. The term "promoter" is sometimes used in the art to generically indicate a promoter region. Many different promoters are known in the art which direct expression of a gene in a certain cell types. Tissue-specific promoters can comprise nucleic acid sequences which cause a greater (or decreased) level of expression in cells of a certain tissue type.

"Operably linked" means that the promoter controls the initiation of expression of the gene. A promoter is operably linked to a sequence of proximal DNA if upon introduction into a host cell the promoter determines the transcription of the proximal DNA sequence(s) into one or more species of RNA. A promoter is operably linked to a DNA sequence if the promoter is capable of initiating transcription of that DNA sequence.

"Prokaryote" refers to all organisms without a true nucleus, including bacteria.

"Eukaryote" refers to organisms and cells that have a true nucleus, including mammalian cells.

"Host" includes prokaryotes and eukaryotes, such as yeast and filamentous fungi, as well as plant and animal cells. The term includes an organism or cell that is the recipient of a replicable expression vehicle.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. Preferred animals include higher eukaryotes such as avians, rodents (e.g., mice, rabbits, rats, chinchillas, guinea pigs, hamsters and the like), and mammals. Preferred mammals include bovine, equine, feline, canine, ovine, caprine, porcine, buffalo, humans, and primates.

A "transgenic animal" is an animal containing one or more cells bearing genetic information received, directly or indirectly, by deliberate genetic manipulation or by inheritance from a manipulated progenitor at a subcellular level, such as by microinjection or infection with a recombinant viral vector (e.g., adenovirus, retrovirus, herpes virus, adeno-associated virus, lentivirus). This introduced DNA molecule may be integrated within a chromosome, or it may be extra-chromosomally replicating DNA.

"Embryonic stem cells" or "ES cells" as used herein are cells or cell lines usually derived from embryos which are pluripotent meaning that they are un-differentiated cells. These cells are also capable of incorporating exogenous DNA by homologous recombination and subsequently developing into any tissue in the body when incorporated into a host embryo. It is possible to isolate pluripotent cells from sources other than embryonic tissue by methods which are well understood in the art.

Embryonic stem cells in mice have enabled researchers to select for transgenic cells and perform gene targeting. This allows more genetic engineering than is possible with other transgenic techniques. For example, mouse ES cells are relatively easy to grow as colonies *in vitro*. The cells can be transfected by standard procedures and transgenic cells clonally selected by antibiotic resistance. See, for example, Doetschman *et al.*, 1994, *Gene transfer in embryonic stem cells*. In Pinkert (Ed.) Transgenic Animal Technology: A Laboratory

Handbook. Academic Press Inc., New York, pp.115-146. Furthermore, the efficiency of this process is such that sufficient transgenic colonies (hundreds to thousands) can be produced to allow a second selection for homologous recombinants. Mouse ES cells can then be combined with a normal host embryo and, because they retain their potency, can develop into all the tissues in the resulting chimeric animal, including the germ cells. The transgenic modification can then be transmitted to subsequent generations.

Methods for deriving embryonic stem (ES) cell lines *in vitro* from early preimplantation mouse embryos are well known. See for example, Evans *et al.*, 1981 *Nature* 29: 154-6 and Martin, 1981, *Proc. Nat. Acad. Sci. USA*, 78: 7634-8. ES cells can be passaged in an undifferentiated state, provided that a feeder layer of fibroblast cells or a differentiation inhibiting source is present.

The term "somatic cell" indicates any animal or human cell which is not a sperm or egg cell or is capable of becoming a sperm or egg cell. The term "germ cell" or "germ-line cell" refers to any cell which is either a sperm or egg cell or is capable of developing into a sperm or egg cell and can, therefore pass its genetic information to offspring. The term "germ cell-line transgenic animal" refers to a transgenic animal in which the genetic information was incorporated in a germ line cell, thereby conferring the ability to transfer the information to offspring. If such offspring in fact possess some or all of that information, then they, too, are transgenic animals.

The genetic alteration of genetic information may be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

"Fragment" of a gene refers to any portion of a gene sequence. A "biologically active fragment" refers to any portion of the gene that retains at least one biological activity of that gene. For example, the fragment can perhaps hybridize to its cognate sequence or is capable of being translated into a polypeptide fragment encoded by the gene from which it is derived.

"Variant" refers to a gene that is substantially similar in structure and biological activity or immunological characteristics to either the entire gene or to a fragment of the gene. Provided that the two genes possess a similar activity, they are considered variant as that term is used herein even if the sequence of encoded amino acid residues is not identical.

Preferentially, as used herein (unless otherwise defined) the variant is one of LRP5, HBM or LRP6. The variant preferably is one that yields an HBM-like phenotype (i.e., enhances bones mass and/or modulates lipid levels). These variants include missense mutations, single nucleotide polymorphisms (SNPs), mutations which result in changes in the amino acid sequence of the protein encoded by the gene or nucleic acid, and combinations thereof, as well as com in the exon domains of the *HBM* gene and mutations in LRP5 or LRP6 which result in an HBM like phenotype.

"Amplification of nucleic acids" refers to methods such as polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods based on the use of Q-beta replicase. These methods are well known in the art and described, for example, in U.S. Patent Nos. 4,683,195 and 4,683,202. Reagents and hardware for conducting PCR are commercially available. Primers useful for amplifying sequences from the HBM region are preferably complementary to, and hybridize specifically to sequences in the HBM region or in regions that flank a target region therein. HBM sequences generated by amplification may be sequenced directly. Alternatively, the amplified sequence(s) may be cloned prior to sequence analysis.

"Antibodies" may refer to polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, that can bind to the HBM proteins and fragments thereof or to nucleic acid sequences from the HBM region, particularly from the HBM locus or a portion thereof. Preferred antibodies also include those capable of binding to LRP5, LRP6 and HBM variants. The term antibody is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Proteins may be prepared synthetically in a protein synthesizer and coupled to a carrier molecule and injected over several months into rabbits. Rabbit sera is tested for immunoreactivity to the HBM protein or fragment. Monoclonal antibodies may be made by injecting mice with the proteins, or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with HBM protein or fragments thereof. Harlow *et al.*, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988) and *Using Antibodies: A Laboratory Manual*, Harlow, Ed and Lane, David (Cold Spring Harbor Press, 1999). These antibodies will be

useful in assays as well as pharmaceuticals. By "antibody" is meant to include but not limited to polyclonal, monoclonal, chimeric, human, humanized, bispecific, multispecific, primatizedTM antibodies.

"HBM protein" refers to a protein that is identical to a *Zmax1* (LRP5) protein except that it contains an alteration of glycine 171 to a valine. An HBM protein is defined for any organism that encodes a *Zmax1* (LRP5) true homolog. For example, a mouse HBM protein refers to the mouse *Zmax1* (LRP5) protein having the glycine 170 to valine substitution.

By "HBM-like" is meant a variant of LRP5, LRP6 or HBM which when expressed in a cell is capable of modulating bone mass, lipid levels, Dkk activity, and/or Wnt activity.

In one embodiment of the present invention, "*HBM* gene" refers to the genomic DNA sequence found in individuals showing the HBM characteristic or phenotype, where the sequence encodes the protein indicated by SEQ ID NO: 4. The *HBM* gene and the *Zmax1* (LRP5) gene are allelic. The protein encoded by the *HBM* gene has the property of causing elevated bone mass, while the protein encoded by the *Zmax1* (LRP5) gene does not. The *HBM* gene and the *Zmax1* (LRP5) gene differ in that the *HBM* gene has a thymine at position 582, while the *Zmax1* gene has a guanine at position 582. The *HBM* gene comprises the nucleic acid sequence shown as SEQ ID NO: 2. The *HBM* gene may also be referred to as an "HBM polymorphism." Other *HBM* genes may further have silent mutations, such as those discussed in Section 3 below.

In alternative embodiments of the present invention, "HBM gene" may also refer to any allelic variant of *Zmax1* (LRP5) or LRP6 which results in the HBM phenotype. Such variants may include alteration from the wild-type protein coding sequence as described herein and/or alteration in expression control sequences of *Zmax1* (LRP5) or contains an amino acid mutation in LRP5 or LRP6, such that the resulting protein produces a phenotype which enhances bone mass and/or modulates lipid levels. A preferred example of such a variant is an alteration of the endogenous *Zmax1* (LRP5) promoter region resulting in increased expression of the *Zmax1* (LRP5) protein.

"Normal," "wild-type," "unaffected," "*Zmax1*", "*Zmax*", "LR3" and "LRP5" all refer to the genomic DNA sequence that encodes the protein indicated by SEQ ID NO: 3. LRP5 has also been referred to LRP7 in mouse. *Zmax1*, LRP5 and *Zmax* may be used

interchangeably throughout the specification and are meant to be the same gene, perhaps only relating to the gene in a different organism. The *Zmax1* gene has a guanine at position 582 in the human sequence. The *Zmax1* gene of human comprises the nucleic acid sequence shown as SEQ ID NO: 1. "Normal," "wild-type," "unaffected", "Zmax1" and "LRP5" also refer to
5 allelic variants of the genomic sequence that encodes proteins that do not contribute to elevated bone mass. The *Zmax1* (*LRP5*) gene is common in the human population, while the *HBM* gene is rare.

"5YWTD+EGF" refers to a repeat unit found in the *Zmax1* (*LRP5*) protein, consisting of five YWTD repeats followed by an EGF repeat.

10 "Bone development" generally refers to any process involved in the change of bone over time, including, for example, normal development, changes that occur during disease states, and changes that occur during aging. This may refer to structural changes in and dynamic rate changes such as growth rates, resorption rates, bone repair rates, and etc. "Bone development disorder" particularly refers to any disorders in bone development including, for
15 example, changes that occur during disease states and changes that occur during aging. Bone development may be progressive or cyclical in nature. Aspects of bone that may change during development include, for example, mineralization, formation of specific anatomical features, and relative or absolute numbers of various cell types.

"Bone modulation" or "modulation of bone formation" refers to the ability to affect
20 any of the physiological processes involved in bone remodeling, as will be appreciated by one skilled in the art, including, for example, bone resorption and appositional bone growth, by, *inter alia*, osteoclastic and osteoblastic activity, and may comprise some or all of bone formation and development as used herein.

Bone is a dynamic tissue that is continually adapting and renewing itself through the
25 removal of old or unnecessary bone by osteoclasts and the rebuilding of new bone by osteoblasts. The nature of the coupling between these processes is responsible both for the modeling of bone during growth as well as the maintenance of adult skeletal integrity through remodeling and repair to meet the everyday needs of mechanical usage. There are a number of diseases of bone that result from an uncoupling of the balance between bone resorption and
30 formation. With aging there is a gradual "physiologic" imbalance in bone turnover, which is

particularly exacerbated in women due to menopausal loss of estrogen support, that leads to a progressive loss of bone. The reduction in bone mass and deterioration in bone architecture results in an increase in bone fragility and susceptibility to spontaneous fractures. For every 10 percent of bone that is lost the risk of fracture doubles. Individuals with bone mineral density (BMD) in the spine or proximal femur 2.5 or more standard deviations below normal peak bone mass are classified as osteoporotic. However, osteopenic individuals with BMD between 1 and 2.5 standard deviations below the norm are clearly at risk of suffering bone loss related disorders.

Bone modulation may be assessed by measuring parameters such as bone mineral density (BMD) and bone mineral content (BMC) by pDXA X-ray methods, bone size, thickness or volume as measured by X-ray, bone formation rates as measured for example by calcien labeling, total, trabecular, and mid-shaft density as measured by pQCT and/or mCT methods, connectivity and other histological parameters as measured by mCT methods, mechanical bending and compressive strengths as preferably measured in femur and vertebrae respectively. Due to the nature of these measurements, each may be more or less appropriate for a given situation as the skilled practitioner will appreciate. Furthermore, parameters and methodologies such as a clinical history of freedom from fracture, bone shape, bone morphology, connectivity, normal histology, fracture repair rates, and other bone quality parameters are known and used in the art. Most preferably, bone quality may be assessed by the compressive strength of vertebra when such a measurement is appropriate. Bone modulation may also be assessed by rates of change in the various parameters. Most preferably, bone modulation is assessed at more than one age.

"Normal bone density" refers to a bone density within two standard deviations of a Z score of 0 in the context of the HBM linkage study. In a general context, the range of normal bone density parameters is determined by routine statistical methods. A normal parameter is within about 1 or 2 standard deviations of the age and sex normalized parameter, preferably about 2 standard deviations. A statistical measure of meaningfulness is the P value which can represent the likelihood that the associated measurement is significantly different from the mean. Significant P values are $P < 0.05$, 0.01, 0.005, and 0.001, preferably at least $P < 0.01$.

"HBM" refers to "high bone mass" although this term may also be expressed in terms

of bone density, mineral content, and size.

The "HBM phenotype" and "HBM-like phenotype" may be characterized by an increase of about 2 or more standard deviations, preferably 2, 2.5, 3, or more standard deviations in 1, 2, 3, 4, 5, or more quantitative parameters of bone modulation, preferably bone density and mineral content and bone strength parameters, above the age and sex norm for that parameter. The HBM phenotype and HBM-like phenotype are characterized by statistically significant increases in at least one parameter, preferably at least 2 parameters, and more preferably at least 3 or more parameters. The HBM phenotype and the HBM-like phenotype may also be characterized by an increase in one or more bone quality parameters and most preferably increasing parameters are not accompanied by a decrease in any bone quality parameters. Most preferably, an increase in bone modulation parameters and/or bone quality measurements is observed at more than one age. The HBM phenotype and HBM-like phenotype also includes changes of lipid levels, Wnt activity and/or Dkk activity.

A "Zmax1 system" or "LRP5 system" refers to a purified protein, cell extract, cell, animal, human or any other composition of matter in which Zmax1 (LRP5) is present in a normal or mutant form.

The terms "isolated" and "purified" refer to a substance altered by hand of man from the natural environment. An isolated peptide may be for example in a substantially pure form or otherwise displaced from its native environment such as by expression in an isolated cell line or transgenic animal. An isolated sequence may for example be a molecule in substantially pure form or displaced from its native environment such that at least one end of said isolated sequence is not contiguous with the sequence it would be contiguous with in nature.

A "surrogate marker" refers to a diagnostic indication, symptom, sign or other feature that can be observed in a cell, tissue, human or animal that is correlated with the *HBM* gene or elevated bone mass or both, but that is easier to measure than bone density. The general concept of a surrogate marker is well accepted in diagnostic medicine.

The present invention encompasses the *Zmax1* (*LRP5*) gene and Zmax1 (LRP5) protein in the forms indicated by SEQ ID NOS: 1 and 3, respectively, and other closely related variants, as well as the adjacent chromosomal regions of Zmax1 (LRP5) necessary for

its accurate expression. In a preferred embodiment, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO: 1.

The present invention further encompasses variants of the *LRP6* gene and its corresponding protein which result in an enhanced bone mass and/or modulate lipids and/or modulate the Wnt signaling pathway.

"Biologically active" refers to those forms of proteins and polypeptides, including conservatively substituted variants, alleles of genes encoding a protein or polypeptide fragments of proteins which retain a biological and/or immunological activity of the wild-type protein or polypeptide. Preferably the activity is one which induces a change in LRP5, LRP6, or Dkk activity, such as inhibiting the interaction of LRP5 or LRP6 or variants thereof with Dkk, or Dkk with another ligand binding partner (*e.g.*, Dkk-1 with a Dkk-1 interacting protein such as those shown in Fig. 31). By biologically active is also meant to include any form which modulates Wnt signaling.

By "modulate" and "regulate" is meant methods, conditions, or agents which increase or decrease the wild-type activity of an enzyme, inhibitor, signal transducer, receptor, transcription activator, co-factor, and the like. This change in activity can be an increase or decrease of mRNA translation, mRNA or DNA transcription, and/or mRNA or protein degradation, which may in turn correspond to an increase or decrease in biological activity.

By "modulated activity" and "regulated activity" is meant any activity, condition, disease or phenotype which is modulated by a biologically active form of a protein. Modulation may be effected by affecting the concentration or subcellular localization of biologically active protein, *i.e.*, by regulating expression or degradation, or by direct agonistic or antagonistic effect as, for example, through inhibition, activation, binding, or release of substrate, modification either chemically or structurally, or by direct or indirect interaction which may involve additional factors.

By "effective amount" or "dose effective amount" or "therapeutically effective amount" is meant an amount of an agent which modulates a biological activity of the polypeptide of the invention.

By "immunologically active" is meant any immunoglobulin protein or fragment thereof which recognizes and binds to an antigen.

By "Dkk" is meant to refer to the nucleic acids and proteins of members of the Dkk (Dickkopf) family. This includes, but is not limited to, Dkk-1, Dkk-2, Dkk-3, Dkk-4, Soggy, and related Dkk proteins. Dkk-1 is a preferred embodiment of the present invention. However, the Dkk proteins have substantial homology and one skilled in the art will appreciate that all of the embodiments of the present invention utilizing Dkk-1 may also be utilized with the other Dkk proteins.

By "Dkk-1" is meant to refer to the Dkk-1 protein and nucleic acids which encode the Dkk-1 protein. Dkk-1 refers to Dickkopf-1, and in *Xenopus* it is related to at least Dkk-2, Dkk-3, and Dkk-4 (see Krupnik *et al.*, 1999 *Gene* 238: 301-313). Dkk-1 was first identified in *Xenopus* (Glinka *et al.*, 1998 *Nature* 391: 357-62). It was recognized as a factor capable of inducing ectopic head formation in the presence of inhibition of the BMP pathway. It was then also found to inhibit the axis-inducing activity of several *Xenopus* Wnt molecules by acting as an extracellular antagonist of Wnt signaling. Mammalian homologs have been found including Dkk-1, Dkk-2, Dkk-3, Dkk-4 and soggy (Fedi *et al.*, 1999 *J. Biol. Chem.* 274: 19465-72; and Krupnick *et al.* 1999). Human Dkk-1 was also referred to as sk (Fedi *et al.* 1999). As used herein, Dkk-1 is meant to include proteins from any species having a Wnt pathway in which Dkk-1 interacts. Particularly preferred are mammalian species (*e.g.*, murine, caprine, canine, bovine, feline, equine, primate, ovine, porcine and the like), with particularly preferred mammals being humans. Nucleic acid sequences encoding Dkk-1 include, but are not limited to human Dkk-1 (GenBank Accession Nos. AH009834, XM_005730, AF261158, AF261157, AF177394, AF127563 and NM_012242), *Mus musculus* dickkopf homolog 1 (GenBank Accession No. NM_010051), and *Danio rerio* dickkopf-1 (GenBank Accession Nos. AF116852 and AB023488). The genomic sequences with exon annotation are GenBank Accession Nos. AF261157 and AF261158. Also contemplated are homologs of these sequences which have Dkk-1 activity in the Wnt pathway. Dkk-1 amino acid sequences include, but are not limited to human dickkopf homolog 1 (GenBank Accession Nos. AAG15544, BAA34651, NP_036374, AAF02674, AAD21087, and XP_005730), *Danio rerio* (zebrafish) dickkopfl (GenBank Accession Nos. BAA82135 and AAD22461) and murine dickkopf-1 (GenBank Accession Nos. O54908 and NP_034181). Variants and homologs of these sequences which possess Dkk-1 activity are

also included when referring to Dkk-1.

By "LRP5 mediated", "LRP6 mediated", and "Dkk mediated" disorder, condition or disease is any abnormal state that involves LRP5, LRP6 and/or Dkk activity. The abnormal state can be induced by environmental exposure or drug administration. Alternatively, the disease or disorder can be due to a genetic defect. Dkk mediated diseases, disorders and conditions include but are not limited to bone mass disorders or conditions and lipid disorders and conditions. For example, bone mass disorders/conditions/diseases, which may be mediated by Dkk, LRP5 and/or LRP6 include but are not limited to age related loss of bone, bone fractures (*e.g.*, hip fracture, Colle's fracture, vertebral crush fractures), chondrodystrophies, drug-induced disorders (*e.g.*, osteoporosis due to administration of glucocorticoids or heparin and osteomalacia due to administration of aluminum hydroxide, anticonvulsants, or glutethimide), high bone turnover, hypercalcemia, hyperostosis, osteoarthritis, osteogenesis imperfecta, osteomalacia, osteomyelitis, osteoporosis, Paget's disease, and rickets.

Lipid disorders/diseases/conditions, which may be mediated by Dkk, LRP5, and/or LRP6 include but are not limited to familial lipoprotein lipase deficiency, familial apoprotein CII deficiency, familial type 3 hyperlipoproteinemia, familial hypercholesterolemia, familial hypertriglyceridemia, multiple lipoprotein-type hyperlipidemia, elevated lipid levels due to dialysis and/or diabetes, and elevated lipid levels of unknown etiologies

The term "recognizes and binds," when used to define interactions of antisense nucleotides or siRNA's (small inhibitory RNA) with a target sequence, means that a particular antisense or small inhibitory RNA (siRNA) sequence is substantially complementary to the target sequence, and thus will specifically bind to a portion of an mRNA encoding polypeptide. As such, typically the sequences will be highly complementary to the mRNA target sequence, and will have no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 base mismatches throughout the sequence. In many instances, it may be desirable for the sequences to be exact matches, *i.e.* be completely complementary to the sequence to which the oligonucleotide specifically binds, and therefore have zero mismatches along the complementary stretch. As such, highly complementary sequences will typically bind quite specifically to the target sequence region of the mRNA and will therefore be highly efficient in reducing, and/or even

inhibiting the translation of the target mRNA sequence into polypeptide product.

Substantially complementary oligonucleotide sequences will be greater than about 80 percent complementary (or '% exact-match') to the corresponding mRNA target sequence to which the oligonucleotide specifically binds, and will, more preferably be greater than about 85 percent complementary to the corresponding mRNA target sequence to which the oligonucleotide specifically binds. In certain aspects, as described above, it will be desirable to have even more substantially complementary oligonucleotide sequences for use in the practice of the invention, and in such instances, the oligonucleotide sequences will be greater than about 90 percent complementary to the corresponding mRNA target sequence to which the oligonucleotide specifically binds, and may in certain embodiments be greater than about 95 percent complementary to the corresponding mRNA target sequence to which the oligonucleotide specifically binds, and even up to and including 96%, 97%, 98%, 99%, and even 100% exact match complementary to the target mRNA to which the designed oligonucleotide specifically binds.

Percent similarity or percent complementary of any of the disclosed sequences may be determined, for example, by comparing sequence information using the GAP computer program, version 6.0, available from the University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (1970). Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for nucleotides, and the weighted comparison matrix of Gribskov and Burgess (1986), (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

By "mimetic" is meant a compound or molecule that performs the same function or behaves similarly to the compound mimicked.

By "reporter element" is meant a polynucleotide that encodes a polypeptide capable of being detected in a screening assays. Examples of polypeptides encoded by reporter elements include, but are not limited to, lacZ, GFP, luciferase, and chloramphenicol acetyltransferase.

2. Introduction

The present invention also encompasses the HBM gene and HBM protein in the forms indicated by SEQ ID NO: 2 and 4, respectively, and other closely related variants, as well as the adjacent chromosomal regions of the HBM gene necessary for its accurate expression. In
5 a preferred embodiment, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO: 2. More preferably, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO: 2, wherein one of the 15 contiguous nucleotides is the thymine at nucleotide 582.

10 The invention also relates to the nucleotide sequence of the *Zmax1* (*LRP5*) gene region, as well as the nucleotide sequence of the HBM region. More particularly, a preferred embodiment are the BAC clones containing segments of the *Zmax1* (*LRP5*) gene region B200E21-H and B527D12-H. A preferred embodiment is the nucleotide sequence of the BAC clones consisting of SEQ ID NOS: 5-12.

15 The invention also concerns the use of the nucleotide sequence to identify DNA probes for the *Zmax1* (*LRP5*) gene and the *HBM* gene, PCR primers to amplify the *Zmax1* (*LRP5*) gene and the *HBM* gene, nucleotide polymorphisms in the *Zmax1* (*LRP5*) gene and the *HBM* gene, and regulatory elements of the *Zmax1* (*LRP5*) gene and the *HBM* gene.

This invention describes the further localization of the chromosomal location of the
20 *Zmax1* (*LRP5*) gene and *HBM* gene on chromosome 11q13.3, between genetic markers D11S987 and SNP_CONTIG033-6, as well as the DNA sequences of the *Zmax1* (*LRP5*) gene and the *HBM* gene. The chromosomal location was refined by the addition of more genetic markers to the mapping panel used to map the gene, and by the extension of the pedigree to include more individuals. The pedigree extension was critical because the new individuals
25 that have been genotyped harbor critical recombination events that narrow the region. To identify genes in the region on 11q13.3, a set of BAC clones containing this chromosomal region was identified. The BAC clones served as a template for genomic DNA sequencing, and also as a reagent for identifying coding sequences by direct cDNA selection. Genomic sequencing and direct cDNA selection were used to characterize more than 1.5 million base
30 pairs of DNA from 11q13.3. The *Zmax1* (*LRP5*) gene was identified within this region and

the *HBM* gene was then discovered after mutational analysis of affected and unaffected individuals.

When a gene has been genetically localized to a specific chromosomal region, the genes in this region can be characterized at the molecular level by a series of steps that include: cloning of the entire region of DNA in a set of overlapping clones (physical mapping), characterization of genes encoded by these clones by a combination of direct cDNA selection, exon trapping and DNA sequencing (gene identification), and identification of mutations in these genes by comparative DNA sequencing of affected and unaffected members of the HBM kindred (mutation analysis).

Physical mapping is accomplished by screening libraries of human DNA cloned in vectors that are propagated in *E. coli* or *S. cerevisiae* using PCR assays designed to amplify unique molecular landmarks in the chromosomal region of interest. To generate a physical map of the HBM candidate region, a library of human DNA cloned in Bacterial Artificial Chromosomes (BACs) was screened with a set of Sequence Tagged Site (STS) markers that had been previously mapped to chromosome 11q12-q13 by the efforts of the Human Genome Project.

STSs are unique molecular landmarks in the human genome that can be assayed by PCR. Through the combined efforts of the Human Genome Project, the location of thousands of STSs on the twenty-two autosomes and two sex chromosomes has been determined. For a positional cloning effort, the physical map is tied to the genetic map because the markers used for genetic mapping can also be used as STSs for physical mapping. By screening a BAC library with a combination of STSs derived from genetic markers, genes, and random DNA fragments, a physical map comprised of overlapping clones representing all of the DNA in a chromosomal region of interest can be assembled.

BACs are cloning vectors for large (80 kilobase to 200 kilobase) segments of human or other DNA that are propagated in *E. coli*. To construct a physical map using BACs, a library of BAC clones is screened so that individual clones harboring the DNA sequence corresponding to a given STS or set of STSs are identified. Throughout most of the human genome, the STS markers are spaced approximately 20 to 50 kilobases apart, so that an individual BAC clone typically contains at least two STS markers. In addition, the BAC

libraries that were screened contain enough cloned DNA to cover the human genome six times over. Therefore, an individual STS typically identifies more than one BAC clone. By screening a six-fold coverage BAC library with a series of STS markers spaced approximately 50 kilobases apart, a physical map consisting of a series of overlapping BAC clones, i.e. BAC contigs, can be assembled for any region of the human genome. This map is closely tied to the genetic map because many of the STS markers used to prepare the physical map are also genetic markers.

When constructing a physical map, it often happens that there are gaps in the STS map of the genome that result in the inability to identify BAC clones that are overlapping in a given location. Typically, the physical map is first constructed from a set of STSs that have been identified through the publicly available literature and World Wide Web resources. The initial map consists of several separate BAC contigs that are separated by gaps of unknown molecular distance. To identify BAC clones that fill these gaps, it is necessary to develop new STS markers from the ends of the clones on either side of the gap. This is done by sequencing the terminal 200 to 300 base pairs of the BACs flanking the gap, and developing a PCR assay to amplify a sequence of 100 or more base pairs. If the terminal sequences are demonstrated to be unique within the human genome, then the new STS can be used to screen the BAC library to identify additional BACs that contain the DNA from the gap in the physical map. To assemble a BAC contig that covers a region the size of the HBM candidate region (2,000,000 or more base pairs), it is often necessary to develop new STS markers from the ends of several clones.

After building a BAC contig, this set of overlapping clones serves as a template for identifying the genes encoded in the chromosomal region. Gene identification can be accomplished by many methods. Three methods are commonly used: (1) a set of BACs selected from the BAC contig to represent the entire chromosomal region can be sequenced, and computational methods can be used to identify all of the genes, (2) the BACs from the BAC contig can be used as a reagent to clone cDNAs corresponding to the genes encoded in the region by a method termed direct cDNA selection, or (3) the BACs from the BAC contig can be used to identify coding sequences by selecting for specific DNA sequence motifs in a procedure called exon trapping. The present invention includes genes identified by the first

two methods.

To sequence the entire BAC contig representing the HBM candidate region, a set of BACs was chosen for subcloning into plasmid vectors and subsequent DNA sequencing of these subclones. Since the DNA cloned in the BACs represents genomic DNA, this sequencing is referred to as genomic sequencing to distinguish it from cDNA sequencing. To initiate the genomic sequencing for a chromosomal region of interest, several non-overlapping BAC clones are chosen. DNA for each BAC clone is prepared, and the clones are sheared into random small fragments which are subsequently cloned into standard plasmid vectors such as pUC18. The plasmid clones are then grown to propagate the smaller fragments, and these are the templates for sequencing. To ensure adequate coverage and sequence quality for the BAC DNA sequence, sufficient plasmid clones are sequenced to yield six-fold coverage of the BAC clone. For example, if the BAC is 100 kilobases long, then phagemids are sequenced to yield 600 kilobases of sequence. Since the BAC DNA was randomly sheared prior to cloning in the phagemid vector, the 600 kilobases of raw DNA sequence can be assembled by computational methods into overlapping DNA sequences termed sequence contigs. For the purposes of initial gene identification by computational methods, six-fold coverage of each BAC is sufficient to yield ten to twenty sequence contigs of 1000 base pairs to 20,000 base pairs.

The sequencing strategy employed in this invention was to initially sequence "seed" BACs from the BAC contig in the HBM candidate region. The sequence of the "seed" BACs was then used to identify minimally overlapping BACs from the contig, and these were subsequently sequenced. In this manner, the entire candidate region was sequenced, with several small sequence gaps left in each BAC. This sequence served as the template for computational gene identification. One method for computational gene identification is to compare the sequence of BAC contig to publicly available databases of cDNA and genomic sequences, *e.g.*, unigene, dbEST, GenBank. These comparisons are typically done using the BLAST family of computer algorithms and programs (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)). The BAC sequence can also be translated into protein sequence, and the protein sequence can be used to search publicly available protein databases, using a version of BLAST designed to analyze protein sequences (Altschul *et al.*, 1997 *Nucl. Acids Res.* 25:

3389-3402). Another method is to use computer algorithms such as MZEF (Zhang, 1997 *Proc. Natl. Acad. Sci. USA*, 94: 565-568) and GRAIL (Uberbacher *et al.*, 1996 *Methods Enzymol.* 266: 259-281), which predict the location of exons in the sequence based on the presence of specific DNA sequence motifs that are common to all exons, as well as the presence of codon usage typical of human protein encoding sequences.

In addition to identifying genes by computational methods, genes were also identified by direct cDNA selection (Del Mastro *et al.*, 1995 *Genome Res.* 5(2): 185-194). In direct cDNA selection, cDNA pools from tissues of interest are prepared, and the BACs from the candidate region are used in a liquid hybridization assay to capture cDNA which base-pairs to coding regions in the BAC. In the methods described herein, the cDNA pools were created from several different tissues by random priming the first strand cDNA from poly-A RNA, synthesizing the second strand cDNA by standard methods, and adding linkers to the ends of the cDNA fragments. The linkers are used to amplify the cDNA pools. The BAC clones are used as a template for *in vitro* DNA synthesis to create a biotin labeled copy of the BAC DNA. The biotin labeled copy of the BAC DNA is then denatured and incubated with an excess of the PCR amplified, linkered cDNA pools which have also been denatured. The BAC DNA and cDNA are allowed to anneal in solution, and heteroduplexes between the BAC and the cDNA are isolated using streptavidin coated magnetic beads. The cDNA which is captured by the BAC is then amplified using primers complimentary to the linker sequences, and the hybridization/selection process is repeated for a second round. After two rounds of direct cDNA selection, the cDNA fragments are cloned, and a library of these direct selected fragments is created.

The cDNA clones isolated by direct selection are analyzed by two methods. Since a pool of BACs from the HBM candidate region is used to provide the genomic DNA sequence, the cDNAs must be mapped to individual BACs. This is accomplished by arraying the BACs in microtiter dishes, and replicating their DNA in high density grids. Individual cDNA clones are then hybridized to the grid to confirm that they have sequence identity to an individual BAC from the set used for direct selection, and to determine the specific identity of that BAC. cDNA clones that are confirmed to correspond to individual BACs are sequenced. To determine whether the cDNA clones isolated by direct selection share

sequence identity or similarity to previously identified genes, the DNA and protein coding sequences are compared to publicly available databases using the BLAST family of programs.

5 The combination of genomic DNA sequence and cDNA sequence provided by BAC sequencing and by direct cDNA selection yields an initial list of putative genes in the region. The genes in the region were all candidates for the HBM locus. To further characterize each gene, Northern blots were performed to determine the size of the transcript corresponding to each gene, and to determine which putative exons were transcribed together to make an individual gene. For Northern blot analysis of each gene, probes were prepared from direct
10 selected cDNA clones or by PCR amplifying specific fragments from genomic DNA or from the BAC encoding the putative gene of interest. The Northern blots gave information on the size of the transcript and the tissues in which it was expressed. For transcripts which were not highly expressed, it was sometimes necessary to perform a reverse transcription PCR assay using RNA from the tissues of interest as a template for the reaction.

15 Gene identification by computational methods and by direct cDNA selection provides unique information about the genes in a region of a chromosome. When genes are identified, then it is possible to examine different individuals for mutations in each gene.

The present invention also encompasses the *HBM* gene and HBM protein in the forms indicated by SEQ ID NO: 2 and 4, respectively, and other closely related variants, as well as
20 the adjacent chromosomal regions of the *HBM* gene necessary for its accurate expression. In a preferred embodiment, the present invention is directed to an isolated nucleic acid sequence of SEQ ID NO: 2, as well as variants thereof. Variants include changes in SEQ ID NO:1 which result in a HBM like phenotype. Examples of such variants are discussed further in Section 3 below and in the examples. These variants preferably have at least about 90%,
25 preferably at least about 95%, or more preferably at least about 98% or more similarity or identity to the nucleic acid sequence of SEQ ID NOS: 1 or 2 or biologically active fragments thereof. Therefore, sequences which are 96%, 97%, and 99% or more similar to SEQ ID NOS: 1 or 2 or biologically active fragments thereof are also contemplated herein.

30 Determination of the degree of variation between a high bone mass (HBM) variant can be performed using BLAST or FASTA or other suitable algorithm using standard default

parameters. Preferably, identity will be determined for coding regions of SEQ ID NOS: 1-2, but can also include non-coding domains. Additionally, alignment programs can be used to identify conserved sequences or potential motifs across different animal species. Alignment programs can also be used to align the nucleic acid and/or protein sequences of related genes and the proteins that they encode. Preferred alignment programs include CLUSTALW, PILEUP and GAP, and would preferably be used with default parameters. For example, such programs can be used to align the sequences of Zmax1 (LRP5), HBM, and LDL receptor-related protein 6 (LRP6) and sequences related thereto.

By a polynucleotide having a nucleotide sequence at least, for example, 90% "similar" to a reference nucleotide sequence encoding a polypeptide, is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to ten point mutations per each 100 nucleotides of the reference nucleotide sequence. These mutations of the reference sequence may occur at any location in SEQ ID NO: 1 or 2 or in the *LRP6* gene. The mutations may be silent.

Another embodiment contemplates that such polynucleotide variants of SEQ ID NO: 1 or 2 comprise nucleic acid sequences which are at least 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 300, 400, or 500 contiguous nucleotides of SEQ ID NO: 1 or 2. More preferably, such polynucleotide variants have a contiguous nucleic acid sequence corresponding with the polymorphism at nucleotide 582 (G→T substitution) of SEQ ID NO: 2 or other variants of SEQ ID NO: 1 or 2, which comprise a mutation which modulates bone mass and/or lipid levels when the polypeptide encoded thereby is administered to a subject. All variants of SEQ ID NO: 1 or 2 contemplated possess the characteristic of encoding a protein or polypeptide which when administered to a subject induces bone modulation. Additional variants which may be responsible for modulating bone mass when administered to a subject may lie within the domain known to contain the HBM polymorphism and which encodes the beta propeller domain (YWTD motifs). Alternatively, other variants of Zmax1 (LRP5) which modulate high bone mass in a subject may be due to mutations in the nucleic acid sequences encoding any of the other conserved domains of Zmax1 (LRP5), such as those set forth in Figure 4 (e.g., the RGD extracellular attachment site, the binding site for LDL and calcium, the cysteine rich growth factor repeats, the ideal PEST region, and the

internalization domain). See Section 3 and the Examples below for additional mutations which may confer enhance bone mass and/or lipid modulation.

HBM polynucleotides and HBM like variants contemplated include those which hybridize under stringent conditions to SEQ ID NO: 2. Hybridization methods are known in the art and include, but are not limited to: (a) washing with 0.1X SSPE (0.62 M NaCl, 0.06 M NaH₂PO₄·H₂O, 0.075 M EDTA, pH 7.4) and 0.1% sodium dodecyl sulfate (SDS) at 50°C; (b) washing with 50% formamide, 5X SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6-8), 0.1% sodium pyrophosphate, 5X Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS and 10% dextran sulfate at 42°C, followed by washing at 42°C in 0.2X SSC and 0.1% SDS; and (c) washing with 0.5 M NaPO₄, 7% SDS at 65°C followed by washing at 60°C in 0.5X SSC and 0.1% SDS. Additional conditions under which HBM variants can be isolated by hybridization to SEQ ID NO: 2 or nucleic acid fragments thereof can be performed by varying the hybridization temperature. High stringency hybridization conditions are those performed at about 20°C below the melting temperature (T_m) of SEQ ID NO: 2 or fragments thereof. Preferred stringency is performed at about 5-10°C below the T_m of SEQ ID NO: 2 or fragments thereof. Additional hybridization conditions can be prepared as described in Chapter 11 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (1989), or as would be known to the artisan of ordinary skill.

Alternatively, mammalian libraries (e.g., equine, primate, caprine, bovine, ovine, feline, porcine, and canine) can be probed using degenerate primers and polymerase chain reaction (PCR) techniques to identify variants of SEQ ID NO: 2 or fragments thereof. Preferably primers are utilized which hybridize under stringent conditions to the open reading frame of SEQ ID NO: 2, or to non-coding portions of the sequence. More preferably, such primers hybridize to conserved domains within SEQ ID NO: 2. For example, conserved domains include those coding for the YWTD beta-propeller domains or other domains, such as those listed in Figure 4. Preferred primers are typically 15 nucleotides in length, but can vary to be at least, about 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40 or 50 (and any range in between) nucleotides in length. Heterologous hybridization is to amplify the target gene or nucleic acid sequence using degenerate PCR primers. Probes for variants of SEQ ID NO: 2

and the polypeptide encoded thereby can be obtained by preparing mixed oligonucleotides of greater than 10, preferably of 15 or more, nucleotides in length representing all possible nucleotide sequences which could encode the corresponding amino acid sequences (e.g., SEQ ID NO: 4 fragments thereof). This method is clearly documented by Gould *et al.*, 1989 *Proc. Natl. Acad. Sci. USA* 86(6): 1934-8.

Another embodiment includes nucleic acids which encode an HBM polypeptide or HBM like variant which is at least about 90% similar to SEQ ID NO: 4 and fragments thereof, and which when administered to a subject modulate bone mass in that subject. HBM variants may have a valine corresponding to position 171 of SEQ ID NO: 4 (Gly to Val substitution) or 170 of the mouse homolog, or an amino acid change elsewhere in propeller 1 or in the protein which results in enhanced bone mass and/or lipid modulation. Other preferred embodiments include high bone mass polypeptides which have at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 500 or more contiguous amino acids of SEQ ID NO: 3 or 4, with a mutation in the sequence resulting in enhanced bone mass and/or modulated lipid levels. Such contemplated contiguous sequences preferably overlap with a polymorphism corresponding to high bone mass, such as valine-171 of SEQ ID NO: 4, other mutations in propeller 1 or is predicted from the model provided in the Examples. Also contemplated are the polynucleotides encoding polypeptides which are at least about 95%, 96%, 97%, 98% and 99% or more similar to SEQ ID NO: 3 or 4 and fragments thereof, wherein these polypeptide contain at least one mutation (e.g., valine-171 or like mutation).

In another embodiment, a synthetic nucleic acid encoding SEQ ID NO: 4 is contemplated wherein the nucleic acid sequence has been conservatively substituted based on the degeneracy of the code such that no amino acids are altered in SEQ ID NO: 4, but perhaps wherein the resulting synthetic polynucleotide encoding said SEQ ID NO: 4 is one that is at least, about 50% similar to SEQ ID NO: 2. Alternatively, SEQ ID NO: 4 may contain any of the silent mutations identified in Section 3 below .

By a polypeptide having an amino acid sequence of at least, for example, 95% "identity" to a reference amino acid sequence of SEQ ID NO: 3 or 4 or fragment thereof is intended that the amino acid sequence of the polypeptide is identical to the reference

sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid sequence of SEQ ID NO: 3 or 4 and has an HBM like phenotype. In other words, to obtain a polypeptide having an amino acid sequence 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Additional HBM polypeptides and nucleic acids which encode said HBM polypeptides are contemplated wherein amino acid residues are conservatively substituted. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie *et al.*, "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science* 247: 1306-10 (1990), wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality. These studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at a certain position of the protein. Numerous phenotypic substitutions are described in Bowie *et al.*, *supra*, and the references cited therein, which are herein incorporated by reference in their entirety.

Preferred substitutions would be in domains which are less conserved across species, and which do not correspond to a structurally or functionally important domain (e.g., a binding site, catalytic site, or beta propeller or other domain described in Figure 4).

Recent studies have indicated that LRP5 participates in the Wnt signal transduction pathway. The Wnt pathway is critical in limb early embryological development. A recently published sketch of the components of Wnt signaling is shown in Fig. 28 (Nusse, 2001

<http://www.stanford.edu/~rnusse/pathways/cell2.html>) (*see also*, Nusse, 2001 *Nature* 411: 255-6; and Mao *et al.*, 2001 *Nature* 411: 321-5).

Briefly summarized, Wnt proteins are secreted proteins which interact with the transmembrane protein Frizzled (Fz). LRP proteins, such as LRP5 and LRP6, are believed to modulate the Wnt signal in a complex with Fz (Tamai *et al.*, 2000 *Nature* 407: 530-5). The Wnt pathway acts intracellularly through the Disheveled protein (Dsh) which in turn inhibits glycogen synthetase kinase-3 (GSK3) from phosphorylating β -catenin. Phosphorylated β -catenin is rapidly degraded following ubiquitination. However, the stabilized β -catenin accumulates and translocates to the nucleus where it acts as a cofactor of the T-cell factor (TCF) transcription activator complex.

The protein dickkopf-1 (Dkk-1) is reported to be an antagonist of Wnt pathway. Dkk-1 is required for head formation in early development (Glinka *et al.*, 1998 *Nature* 391: 357-62). Dkk-1 and its function in the Wnt pathway are described in *e.g.*, Krupnik, *et al.*, 1999 *Gene* 238: 301-13; Fedi *et al.*, 1999 *J. Biol. Chem.* 274: 19465-72; *see also* for Dkk-1 and the Wnt pathway, Wu *et al.*, 2000 *Curr. Biol.* 10: 1611-4; Shinya *et al.*, 2000 *Mech. Dev.* 98: 3-17; Mukhopadhyay *et al.*, 2001 *Dev. Cell.* 1: 423-434; and in PCT Patent Application No. WO 00/52047, and in references cited in each. It has been known that Dkk-1 acts upstream of Dsh, however the nature of the mechanism of inhibition by Dkk-1 is just beginning to be elucidated. Dkk-1 is expressed in the mouse embryonic limb bud and its disruption results in abnormal limb morphogenesis, among other developmental defects (Gotewold *et al.*, 1999 *Mech. Dev.* 89:151-3; and Mukhopadhyay *et al.*, 2001 *Dev. Cell.* 1: 423-34).

Related U.S. Serial No. 60/291,311 (herein incorporated by reference in its entirety) disclosed a novel interaction between Dkk-1 (GenBank Accession No. XM 005730) and LRP5. The interaction between Dkk-1 and LRP5 was discovered by a yeast two hybrid (Y2H) screen for proteins which interact with the ligand binding domain of LRP5. The two-hybrid screen is a common procedure in the art, which is described, for example, by Gietz *et al.*, 1997 *Mol. Cell. Biochem.* 172: 67-79; Young, 1998 *Biol. Reprod.* 58: 302-11; Brent *et al.*, 1997 *Ann. Rev. Genet.* 31: 663-704; and Lu *et al.*, eds., Yeast Hybrid Technologies, Eaton Publishing, Natick MA, (2000). More recently, other studies confirm that Dkk-1 is a binding partner for LRP and modulates the Wnt pathway via direct binding with LRP (R.

Nusse, 2001 *Nature* 411: 255-6; Bafico *et al.*, 2001 *Nat. Cell Biol.* 3: 683-6; Semenov, 2001 *Curr. Biol.* 11: 951-61; Mao, 2001 *Nature* 411: 321-5 (2001); Zorn, 2001 *Curr. Biol.* 11: R592-5); and Li *et al.*, 2002 *J. Biol Chem.* 277: 5977-81).

Mao and colleagues (2001) identified Dkk-1 as a ligand for LRP6. Mao *et al.* suggest
5 that Dkk-1 and LRP6 interact antagonistically where Dkk proteins inhibit the Wnt coreceptor
functions of LRP6. Using co-immunoprecipitation, the group verified that the Dkk-1/LRP6
interaction was direct. Dkk-2 was also found to directly bind LRP6. Contrary to data
contained in provisional application 60/291,311, Mao *et al.* report that no interaction was
detected between any Dkk protein and LRP5, as well as no interaction with LDLR, VLDLR,
10 ApoER, or LRP). Additionally, Mao *et al.* demonstrated that LRP6 can titrate Dkk-1's
effects of inhibiting Wnt signaling using the commercial TCF-luciferase reporter gene assay
(TOPFLASH). A similar conclusion was drawn from analogous studies in *Xenopus* embryos.
Deletion analyses of LRP6 functional domains revealed that EGF repeats (beta-propellers) 3
and 4 were necessary for Dkk-1 binding and that the ligand binding domains of LRP6 had no
15 effect on Dkk-1 binding. The findings of Mao *et al.* contrast with data obtained by the
present inventors indication that the ligand binding domains of LRP5 were necessary and
sufficient for Dkk-1 binding in yeast. Using classical biochemical ligand-receptor studies,
Mao *et al.* determined a $K_d=0.34$ nM for Dkk-1/LRP6 and a $K_d=0.73$ nM for Dkk-2/LRP6.

Semenov *et al.* (2001) verified the Mao group's results and confirmed by
20 coimmunoprecipitation that Dkk-1 does not directly bind to Wnt or Frizzled but rather
interacts with LRP6. Their Scatchard analyses found a $K_d=0.5$ nM for Dkk-1/LRP6.
Semenov *et al.* also demonstrated that Dkk-1 could abolish an LRP5/Frizzled8 complex
implying that Dkk-1 can also repress Wnt signaling via interactions with LRP5. A Dkk-1
mutant where cysteine 220 was changed to alanine abolished LRP6 binding and was unable
25 to repress Wnt signaling. Studies in *Xenopus* embryos confirmed the results and revealed a
functional consequence of Dkk-1/LRP6: repression of Wnt signaling. Their *Xenopus* work
also suggested that LRP6/Dkk-1 may be specific for the canonical, β -catenin-mediated, Wnt
pathways as opposed to the Wnt Planar Cell Polarity pathway.

Bafico *et al.* (2001) employed a ^{125}I -labeled Dkk-1 molecule to identify LRP6 as its
30 sole membrane receptor with a $K_d=0.39$ nM. Again, the functional consequences of the Dkk-

1/LRP6 interaction was a repression of the canonical Wnt signaling even when Dkk-1 was added at extremely low concentrations (30 pM).

Not wishing to be bound by theory, it is believed that the present invention provides an explanation for the mechanism of Dkk-1 inhibition of the Wnt pathway and provides a mechanism whereby the Wnt pathway may be modulated. The present application and related U.S. Serial No. 60/291,311 (which is herein incorporated by reference in its entirety) describe Dkk-1/LRP5 interactions and demonstrate that the interaction between LRP5/LRP6/HBM, and Dkk can be used in a method as an intervention point in the Wnt pathway for an anabolic bone therapeutic or a modulator of lipid metabolism.

As detailed in the Examples, Dkk-1 is able to repress LRP5-mediated Wnt signaling but not HBM-mediated Wnt signaling. This observation is of particular interest because the HBM mutation in LRP5 is a gain of function or activation mutation. That is, Wnt signaling, via the canonical pathway, is enhanced with HBM versus LRP5. Thus, other HBM-like mutations would also function similarly. The present data suggest the mechanism of this functional activation: the inability of Dkk-1 to repress HBM-mediated Wnt signaling. Further investigations of other Wnt or Dkk family members show differential activities in the canonical Wnt pathway that demonstrate the complexity and variability in Wnt signaling that can be achieved depending on the LRP/Dkk/Wnt/Frizzled repertoire that is expressed in a particular cell or tissue. This may attest to the apparent bone specificity of the HBM or HBM like phenotypes in humans and in the HBM transgenic animals.

Furthermore, the present data reveal the importance and functional consequence for the potential structural perturbation of the first beta-propeller domain of LRP5. Our data identified the ligand binding domain of LRP5 as the interacting region with Dkk-1 while the Mao *et al.* publication demonstrated the functional role of propellers 3 and 4 in their LRP6/Dkk-1 studies. In the present invention, we implicate the first beta propeller domain, via the HBM mutation at residue 171, as having a functional consequence in the Dkk-1-mediated Wnt pathway. The involvement of position 171 of propeller 1 may be direct or indirect with Dkk-1. Direct involvement could arise from perturbations of the 3-dimensional structure of the HBM extracellular domain that render Dkk-1 unable to bind. Alternatively, residue 171 of propeller 1 may directly interact with Dkk-1; however, by itself, it is

insufficient to bind and requires other LRP5 domains. Potential indirect candidate molecules may be among the proteins identified the Dkk-1 yeast-two-hybrid experiments.

It may be that the disruption of Dkk activity is not necessarily mediated by enhancing or preventing the binding of Dkk to LRP5/LRP6/HBM. More than one mechanism may be involved. Indeed, the inventors have observed that Dkk-1 binds LRP5, LRP6, and HBM. It is able to effectively inhibit LRP6, and to a slightly lesser extent, LRP5 activity. Further, has been observed that different members of the Dkk family differentially affect LRP5/LRP6/HBM activity. For example, Dkk-1 inhibits LRP5/LRP6/HBM activity while another Dkk may enhance LRP5/LRP6/HBM activity. An endpoint to consider is the modulation of the LRP5/LRP6/HBM activity, not simply binding.

The present disclosure shows that targeting the modulation of the Dkk-1/LRP5 interaction is a therapeutic intervention point for an HBM or HBM-like mimetic agent. A therapeutic agent of the invention may be a small molecule, peptide or nucleic acid aptamer, antibody, or other peptide/protein, etc. Methods of reducing Dkk-1 expression may also be therapeutic using methodologies such as: RNA interference (i.e., siRNAs), small hairpin RNAs (shRNAs), antisense oligonucleotides, morpholino oligonucleotides, PNAs, antibodies to Dkk-1 or Dkk-1 interacting proteins, decoy or scavenger LRP5 or LRP6 receptors, and knockdown of Dkk-1 or Dkk-1 interactor transcription. For discussion of small hairpin RNAs, see Yu *et al.*, 2002 *Proc. Natl. Acad. Sci. USA* 99: 6047-52; Tuschl, 2002 *Nature Biotech.* 20: 446-8; Lee *et al.*, 2002 *Nature Biotech.* 19: 500-5; Paddison *et al.*, 2002 *Genes & Devel.* 16: 948-58; and Brummelkamp *et al.*, 2002 *Science* 296: 550-3.

In an embodiment of the present invention, the activity of Dkk-1 or the activity of a Dkk-1 interacting protein may be modulated for example by binding with a peptide aptamer of the present invention. In another embodiment, LRP5 activity may be modulated by a reagent provided by the present invention (*e.g.*, a peptide aptamer). In another embodiment, the Dkk-1/LRP5 interaction may be modulated by a reagent of the present invention (*e.g.*, a Dkk-1 interacting protein such as those identified in Fig. 31). In another embodiment, the Wnt signal transduction pathway may be modulated by use of one or more of the above methods. In a preferred embodiment of the present invention, the Dkk-1 mediated activity of the Wnt pathway may be specifically modulated by one or more of the above methods. In

another preferred embodiment of the present invention, the Wnt signal transduction pathway may be stimulated by down-regulating Dkk-1 interacting protein activity; such down-regulation could, for example, yield greater LRP5 activity. In a more preferred embodiment, by stimulating LRP5 activity, bone mass regulation may be stimulated to restore or maintain a more optimal level. In another preferred embodiment, by stimulating LRP5 activity, lipid metabolism may be stimulated to restore or maintain a more optimal level. Alternative embodiments provide methods for screening candidate drugs and therapies directed to correction of bone mass disorders or lipid metabolism disorders. And, preferred embodiments of the present invention provide drugs and therapies developed by the use of the reagents and/or methods of the present invention. One skilled in the art will understand that the present invention provides important research tools to develop an effective model of osteoporosis, to increase understanding of bone mass and lipid modulation, and to modulate bone mass and lipid metabolism. For a more detailed description of Dkk-1 and Dkk-1 interacting protein modulation, please refer to U.S. Serial No. 60/361,293, which is herein incorporated by reference in its entirety.

3. Alternative Variants of LRP5/LRP6 Having HBM Activity

A structural model of the LRP5/Zmax1 first beta-propeller module was generated based on a model prediction in Springer *et al.*, 1998 *J. Mol. Biol.* 283: 837-862. Based on the model, certain amino acid residues were identified as important variants of LRP5/HBM/Zmax1. The model and modifications thereof are discussed in more detail in the Examples. The following three categories provide examples of such variants:

The shape of the beta-propeller resembles a disk with inward-sloping sides and a hole down the middle. Residue 171 is in a loop on the outer or top surface of the domain in blade 4 of propeller module 1. Thus, variants comprising changed residues in structurally equivalent positions in other blades; as well as residues that are slightly more interior to the binding pocket, but still accessible to the surface, are important embodiments of the present invention for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention. The following Table contains examples of such variants:

TABLE 1

Variant	Effect of Mutation
A214V	a position equivalent to 171 in blade 5 of propeller 1; alanine is not conserved in other propellers
E128V	a position equivalent to 171 in blade 3 of propeller 1; glutamate is not conserved in other propellers
5 A65V	a position equivalent to 171 in blade 2 of propeller 1; alanine is conserved in propellers 1-3 but not 4
G199V	an accessible interior position in blade 5 of propeller 1; glycine is conserved in propellers 1-3 but not 4
M282V	accessible interior position in blade 1 of propeller 1; methionine is conserved in propellers 1-3 but not 4

These mutations were further analyzed based on a more sophisticated model, as discussed and described in Example 11 below.

LRP5/Zmax1 has four beta-propeller structures; the first three beta-propeller modules conserve a glycine in the position corresponding to residue 171 in human LRP5/Zmax1. Therefore, variants bearing a valine in the equivalent positions in the other propellers are important embodiments of the present invention. The following variants are useful for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention: G479V, G781V, and Q1087V of SEQ ID NO: 3, which demonstrate that propeller 1 is an important determinant of an HBM or HBM-like effect.

The G171V HBM polymorphism (SEQ ID NO: 4) results in “occupied space” of the beta-propeller 1, with the side-chain from the valine residue sticking out into an open binding pocket and potentially altering a ligand/protein interaction. The glycine residue is conserved in LRP5/Zmax1 propellers 1, 2 and 3 but is a glutamine in propeller 4. Therefore, the following variants of LRP5/HBM are important embodiments of the present invention for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention:

G171K: introduces a charged side-chain

G171F: introduces a ringed side-chain

G171I: introduces a branched side-chain

G171Q: introduces the propeller 4 residue

These substitutions along with substitutions in other regions of propeller 1 of SEQ ID NO: 3 (i.e., A214V and M282V) have been shown to produce an HBM-like effect by TCF assay (Figs. 29 and 30). Thus these substitutions in other propeller 1 domains would similarly have an expectation of producing an HBM-like phenotype readily assayable by TCF assay.

Furthermore, LRP6 is the closest homolog of LRP5/Zmax1. LRP6 has a beta-propeller structure predicted to be similar, if not identical to Zmax1. The position corresponding to glycine 171 of human LRP5/Zmax1 is glycine 158 of human LRP6. Thus, corresponding variants of LRP6 are an important embodiment of the present invention for the study of the specificity of LRP5/Zmax1 versus its related family member, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention. Specifically, for example, a glycine to valine substitution at the structurally equivalent position, residue 158, of human LRP6 and similar variants of other species' LRP6 homologs represent important research tools.

Site-directed mutants of LRP5 were generated in the full-length human LRP5 cDNA using the QuikChange XL-Site-Directed Mutagenesis Kit (catalog #200516, Stratagene, La Jolla, CA) following the manufacturer's protocol. The mutant sequences were introduced using complementary synthetic oligonucleotides:

Mutation

Complementary oligos

A65V:	5'-TGGTCAGCGGCCTGGAGGATGTGGCCGCAGTGGACTTCC-3' 5'-GGAAGTCCACTGCGGCCACATCCTCCAGGCCGCTGACCA-3'
E128V	5'-AAGCTGTACTGGACGGACTCAGTGACCAACCGCATCGAGG-3' 5'-CCTCGATGCGGTTGGTCACTGAGTCCGTCCAGTACAGCTT-3'
G171K	5'-ATGTACTGGACAGACTGGAAGGAGACGCCCCGGATTGAGCG-3' 5'-CGCTCAATCCGGGGCGTCTCCTTCCAGTCTGTCCAGTACAT-3'
G171F	5'-ATGTACTGGACAGACTGGTTTGAGACGCCCCGGATTGAGCG-3' 5'-CGCTCAATCCGGGGCGTCTCAAACAGTCTGTCCAGTACAT-3'
G171I	5'-ATGTACTGGACAGACTGGATTGAGACGCCCCGGATTGAGCG-3' 5'-CGCTCAATCCGGGGCGTCTCAATCCAGTCTGTCCAGTACAT-3'

<u>Mutation</u>	<u>Complementary oligos</u>
G171Q	5'-ATGTACTGGACAGACTGGCAGGAGACGCCCCGGATTGAGCG-3' 5'-CGCTCAATCCGGGGCGTCTCCTGCCAGTCTGTCCAGTACAT-3'
G199V	5'-CGGACATTTACTGGCCCAATGTACTGACCATCGACCTGGAGG-3' 5'-CCTCCAGGTCGATGGTCAGTACATTGGGCCAGTAAATGTCCG-3'
A214V	5'-AGCTCTACTGGGCTGACGTCAAGCTCAGCTTCATCCACCG-3' 5'-CGGTGGATGAAGCTGAGCTTGACGTCAGCCCAGTAGAGCT-3'
M282V	5'-GAGTGCCCTCTACTCACCCGTGGACATCCAGGTGCTGAGCC-3' 5'-GGCTCAGCACCTGGATGTCCACGGGTGAGTAGAGGGCACTC-3'
5 G479V	5'-CATGTACTGGACAGACTGGGTAGAGAACCCTAAAATCGAGTGTGC-3' 5'-GCACACTCGATTTTAGGGTCTCTACCCAGTCTGTCCAGTACATG-3'
G781V	5'-CATCTACTGGACCGAGTGGGTCCGCAAGCCGAGGATCGTGCG-3' 5'-CGCACGATCCTCGGCTTGCCGACCCACTCGGTCCAGTAGATG-3'
Q1087V	5'-GTACTTCACCAACATGGTGGACCGGGCAGCCAAGATCGAACG-3' 5'-CGTTCGATCTTGGCTGCCCCGGTCCACCATGTTGGTGAAGTAC-3'
G158V of LRP6	5'-GTACTGGACAGACTGGGTAGAAAGTGCCAAAGATAGAACGTGC-3' 5'-GCACGTTCTATCTTTGGCACTTCTACCCAGTCTGTCCAGTAC-3'

All constructs were sequence verified to ensure that only the engineered modification was present in the gene. Once verified, each variant was functionally evaluated in the TCF-luciferase assay in U2OS cells (essentially as described in Example 6. Other functional evaluations could also be performed, such as the *Xenopus* embryo assay (essentially as described in Example 5), or other assays to evaluate Wnt signaling, Dkk modulation, or anabolic bone effect. Binding of these mutants to Dkk, LRP-interacting proteins, Dkk-interacting proteins, or peptide aptamers to any of the preceding could also be investigated in a variety of ways such as in a two-hybrid system (such as in yeast as described in this application), or other methods.

Figure 29 shows the effects of the G171F mutation in propeller 1 of LRP5. This mutation is at the same position as HBM's G171V substitution. Expression of G171F results in an HBM effect. That is, in the presence of Wnt, G171F is able to activate the TCF-luciferase reporter construct. In fact, it may activate the reporter to a greater extent than either LRP5 or HBM. Furthermore, in the presence of Dkk1 and Wnt1, G171F is less susceptible than LRP5 to modulation by Dkk. These data exemplify that the G171F variant

modulates Wnt signaling in a manner similar to HBM. In addition, this data confirms that HBM's valine residue at 171 is not the only modification at 171 that can result in an HBM effect. Together these data support an important role for LRP5 propeller 1 in modulating Wnt pathway activity; in responding to Dkk modulation; and, in the ability to generate an HBM effect.

Figure 30 shows the effects of the M282V mutation in propeller 1 of LRP5. M282 expression results in an HBM-effect. That is, in the presence of Wnt, M282 is able to activate the TCF-luciferase reporter construct. Furthermore, in the presence of Dkk1 and Wnt1, M282V is less susceptible than LRP5 to modulation by Dkk. These data show that the M282V variant modulates Wnt signaling in a manner similar to HBM. In addition, this data confirms that modifications of other residues in propeller 1 of LRP5 can result in an HBM effect.

These data support an "occupied space" model of the HBM mutation in propeller 1 and show that multiple mutations of propeller 1 are capable of generating an HBM effect; the original G171V HBM mutation is not unique in this ability. Moreover, various perturbations in propeller 1 can modulate Dkk activity.

These data illustrate the molecular mechanism of Dkk modulation of LRP signaling. Using the methods disclosed herein and in U.S. Application 60/290,071, generation of a comprehensive mutant panel will reveal residues in LRP that function in Dkk modulation of Wnt signaling. Such variants of LRP5 and LRP6 that modulate Dkk activity and the residues which distinguish them from LRP5 and LRP6 are points for therapeutic intervention by small molecule compound, antibody, peptide aptamer, or other agents. Furthermore, models of each HBM-effect mutation/polymorphism may be used in rational drug design of an HBM mimetic agent.

These and the examples provided *infra* are only a few illustrative examples presented to better describe the present invention. Variants of LRP5 which have demonstrated HBM activity in assays include A65V, G171I, G171V (HBM), G171F, M282V, G171K, G171Q and A214V. Clearly, other variants may be contemplated within the scope of the present invention. Furthermore, wherever HBM is recited in the methods of the invention, it should be understood that any such alternative variant of LRP5 or LRP6 which demonstrates HBM

like biological activity is also encompassed by those claims.

Additional mutations may also result in conformational changes such as those described above and in the Examples below. These mutations may also result in a HBM-like phenotype when expressed. The following mutations have been identified in Table 2 based on familial genetics in the exons of Zmax1 (LRP5).

TABLE 2: Mutations in the Exons of LRP5

Exon	Location (ATG as nt 1)	Reference	Change	Met as AA1	AA Change
2	249	CGAGGAGGCCATC	C/T	S83	None
2	266	AGACCTACCT	A/G	Q89	G→R
3	512	GTGAGACGCC	G/T	G171	G→V
6	1199	CGTACCTGGACGGG	C/T	A400	A→V
8	1647	CGGGTTCACGCTGC	C/T	F549	None
9	1932	GGCCTTCTTGGTCT	G/A	E644	None
9	1999	GTGGCCATCCCGCT	G/A	V667	V→M
10	2220	CAGAATCGAAGTG	C/T	N740	None
14	3107	GGACACTGTT	G/A	R1036	R→Q
15	3297	CGGCACCGAGCG	C/T	D1099	None
15	3357	AGACAACACACTG	A/G	V1119	None
16	3564	GACTCGCATC	G/A	R1188	None
18	3989	CGGACTGTGA	C/T	A1330	A→V
20	4137	CAGCCCGGCCAC	C/T	D1379	None
20	4248	GGGGGCCAACGG	G/A	A1416	None
20	4565	CGGCCACTGC	C/T	P1522	P→L
23	4635	CGACGTGTGTGACA	C/T	T1545	None

By, for example "C/T" is meant there is a C to T mutation.

Additional mutations have also been identified in the introns, which may result in splice variants are provided in Table 3 below.

TABLE 3: Mutations in the Introns of LRP5

	Exon	Nucleotide Position with respect to Exon	Sequence with SNP Location Underlined	Nucleotide
	2	+53	CTCTCTCTCGAATT	C/T
	5	-4	TCAGTCCACACTCG	T/C
5	5	+8	GGGCGGGGGCTGGG	G/A
	7	-50	GCAGAGACCAGAC	G/A
	8	-118	TGCTCTTGGGCATT	T/G
	9	-131	TGGGGGTGAGTCCT	T/C
	10	+6	IGTTTGCCTGTCCC	T/C
10	11	-173	ATGTGTGTGGCAG	A/G
	11	-152	GGTCTCGCCCTTC	G/A
	11	-49	CGGTGAGAGCAGAC	C/T
	11	+37	CGGGGCAGCCGGG	C/T
	11	+78	GTACCCTGTGGCCT	G/A
15	12	+80	CTCATCTGGGGTTC	C/G
	12	+141	ATGATGCTACCTGG	A/G
	15	-166	CGGGAATTTGGAGA	C/T
	15	-149	TTGTTCAACTAGTA	T/C
	15	-52	TCCGAGGAGACGC	T/G
20	17	-213	TTGTTTCCGGCATC	T/C
	17	-82	CATTGCCCCCTA	C/T
	18	-72	GCCCAGTCAC	G/A
	18	-63	CGCCATTGCC	C/T
	18	-30	GTGTGATGTT	G/A
25	18	+23	TGATCTGGAGGAGG	T/C
	18	+47	GTCTGGGCAGCTTT	G/C
	18	+54	CACCGTCAGTGCT	C/A

Exon	Nucleotide Position with respect to Exon	Sequence with SNP Location Underlined	Nucleotide
22	-118	GGCACCTGCC	G/A

These splice variants may also produce an altered phenotype capable of conferring a HBM-like effect.

5 These mutations were identified using 80 ng of genomic DNA, which was PCR amplified with the primers indicated in Table 4 below with M13F (TGTAACACGACGGCCAGT) attached to the 5' end of the Forward primer or M13R (AGGAAACAGCTATGACCAT) attached to the 5' end of the Reverse primer. 4 μ l of the PCR reaction was diluted to a final volume of 100 μ l with water. Sequencing was performed
10 using ET and M13F and M13R primers or by the ABI-PRISM® Big-Dye™ method of Applied Biosystems using the indicated nested sequencing primers. The sequences are assembled on consed with the appropriate reference sequence.

The PCR mixes used are as follows

Promega: 50 mM KCl, 10 mM Tris-HCl, pH 9.0, 0.1% Triton X-100, 1.5 mM MgCl₂

15 Invitrogen D: 60 mM Tris-HCl, pH 8.5, 15 mM (NH₄)₂SO₄, 3.5 mM MgCl₂

 Invitrogen J: 60 mM Tris-HCL, pH 9.5, 15 mM (NH₄)₂SO₄, 2.0 mM MgCl₂

 InvitrogenM: 60 mM Tris-HCl, pH 10.0, 15 mM (NH₄)₂SO₄, 1.5 mM MgCl₂

20 To all the PCT mixes, 120 μ M of all 4 dNTPs was added, 0.4 μ M of the forward and reverse primers, 80 ng genomic DNA, 1 U of AmpliTaq® DNA polymerase, and 1.1 U of TaqStart antibody (Clontech). The PCR reaction is than run as follows: 94°C, 2 min; (94°C, 30 sec; X-anneal-temp, 30 sec; 72°C, 2 min) for 35 cycles and 72°C for 3 min.

TABLE 4: Forward and Reverse Primers

Exon	F-primer(PCR)	R-primer(PCR)	anneal		Prod.	Size	PCR buffer	F-primer(seq)	R-primer(seq)
			temp	cycle					
1	GAGACGGCGCGGGCTTC	CGCCCCAACTCGTCTCCCAAC	62	2-cy/de-68	429	dmso	TCCGCGCGGCCACGCTC	TGCCAAGTCGCTTCCG	
2	AAGGAACTGGAGGCTTGG	CAGAGTCACACCCCTTTTTC	62		670	promega	GGCATGGGCGAGGGCAGT	TGAAAAACAACCTTGGGCTC	
3	CCAAGTTCGAGAAGTCC	AATACCTGAAACCATACCTG	58		523	promega	TGCATTCTCAGGGGCC	TTGGTTATTTCCGATGGG	
4	GGCGTAGTGGTGGGCATCAG	CCAGGCCAGCCACACACCTC	62		680	dmso	ACTGTCTGGGGACCCCTC	ATTGCAGCAGGTACCCC	
5	GGGAAATTGCAGGCCGCTG	CTGAGGACCAAGGCGGAGAG	58		633	promega	AGGCTGAGGGCCCCATG	CAGGATTGACCTCCTGG	
6	CACCTAACATCACCAGCC	GATGCAAGACAGTGTCCC	62		672	promega	CCTGGCTGAGTATTTC	TCAATCTCCTCTCGCC	
7	ACCCGACATTTACGAGCAC	AAATAGCAGAGCACAGGCAC	58		484	promega+10% dmso	CGACATTTACTGACACCA	CCATGGTGCTCGCCA	
8	TCTCAACAACAACAACAACAAA	TCCTTGGCAGATACTGTAC	60		638	promega	ACAACAACAAGCGTCA	CTTTCTGTCTGCTGCCCT	
9	TGTGTGGCGGGAATAAG	TTGAGGCGAGAACACAGAG	60		648	promega	CTGTGCACATTGGAGCT	CAAGGTTTCCCATAAAGG	
10	ATGTCCTACAAAACACGCTG	CTAATCACTGAGGGCCAC	60		744	promega	GTCTGGCCTGGCGTGG	ACGGACAGCCTGCCACC	
11	TCACCTTAGAGTGAGCC	AGCCTCTCCGACATAAC	60		677	promega	ACTGTGGGAATTCAGGG	TGCAGCAAAAGGCACCCA	
12-1	GAACACGACAAAAGCCC	TGTCCATCAGCAAGTCCAG	60		677	promega	ATGAGGGCGGCCCATGTG	GCGGTTCCGGCGCGCTAG	
12-2	ACAGCGATTATCTACTGGAC	ATGAAAGCACTCAGCAC	60		622	promega	AGACTGGAATCTGCACAG	TGTGCTGGCAGTATGAG	
13	AGAATGAATTTCCCATAGCG	AGACAAAAGTCTGTGGGGTC	52		676	Invitrogen D	TCCCGTGGACCTCCAGC	ACAACGGGGAAACCCAG	
14	GAGAGACCCACACCAATAC	CAGGTGGAAGTCTCCCCAG	63		571	promega	TGGGATTTGACTTTTCAGG	CCTGTGAGAGGCTGGCA	
15	CAGTTGGATTAGGGCCTACC	CAGCTGTCAGTGTGAGGACAA	60		656	promega	CTCACCCATTGTGTCG	AACAGCTACACACAAAG	
16	AACGTATGGCCTTCATCCC	TCTTAAACGCTGTTCGTGGT	60		699	promega	GAGCCAGCCCGAGGTGG	GATTTGTTCTGCGGCAAAAG	
17	AGACTGATGTTATGGGCACAG	ATTGTGAGATGCAGGAACG	60		662	promega	TAGAGACTGGTGCAGAC	CAGTACTTAGAGGAAATTC	
18	ATTCTCCAGCCTCTCTCTG	GCAGGAAGGTGAGATAGACC	58		722	promega	GGTTGGGCTGGAGGTG	TCACAGTCTGCCTTCAAG	
19	GGGGTCTGTTACTCTCTGCAT	CAGCTCACAGTCTGCTCTCT	60		645	promega	GGCGTAGACCTCCCCAC	CCGCTGCCCTGGGAAAG	
20	ACGTTACCCTGAGGTTGGC	AGGCCTCTGTGTTGAAGGATT	60		640	promega	AGGCTCCCGAGGCCTAG	CTGATGCCAAGCAGAGG	
21	TTGGAGGAGGTACCATGTGTC	GGTGGATTGGGTGAGATTTT	55		643	Invitrogen J	TCCCGTTTTCACAGATGAG	CAGATTATCCACAATCAAC	
22	TTCTGCTGATTTCTGAACCC	TTGCCTTTCACTGAGATGAC	58		558	promega+10%dmso	AACATGCACTGCCCGCT	CTTCGGGGCAGGTGGCT	
23	GGCCTGCATCTTCTGGAGC	CATTCCCTCCACGGGGAC	62		683	promega	TGGCCAGTGGACAGGCC	ACACAACCTCAAATGCACAC	

4. Genotyping of Microsatellite Markers

To narrow the genetic interval to a region smaller than that originally reported by Johnson *et al.*, 1997 *Am. J. Hum. Genet.*, 60:1326-1332, additional microsatellite markers on chromosome 11q12-13 were typed. The new markers included: D11S4191, D11S1883,
5 D11S1785, D11S4113, D11S4136, D11S4139, (Dib *et al.*, 1996 *Nature*, 380:152-154), FGF3 (Polymeropoulos *et al.*, 1990 *Nucl. Acid Res.*, 18: 7468), as well as GTC_HBM_Marker_1, GTC_HBM_Marker_2, GTC_HBM_Marker_3, GTC_HBM_Marker_4, GTC_HBM_Marker_5, GTC_HBM_Marker_6, and GTC_HBM_Marker_7 (See Fig. 2).

Blood (20 ml) was drawn into lavender cap (EDTA containing) tubes by a certified
10 phlebotomist. The blood was stored refrigerated until DNA extraction. DNA has been extracted from blood stored for up to 7 days in the refrigerator without reduction in the quality or quantity of yield. For those subjects that have blood drawn at distant sites, a shipping protocol was successfully used on more than a dozen occasions. Blood samples were shipped by overnight express in a styrofoam container with freezer packs to provide
15 cooling. Lavender cap tubes were placed on individual plastic shipping tubes and then into "zip-lock" biohazard bags. When the samples arrived the next day, they were immediately processed to extract DNA.

The DNA extraction procedure used a kit purchased from Gentra Systems, Inc. (Minneapolis, Minnesota). Briefly, the procedure involved adding 3 volumes of a red blood
20 cell lysis buffer to the whole blood. After incubations for 10 minutes at room temperature, the solution was centrifuged in a Beckman tabletop centrifuge at 2,000 X g for 10 minutes. The white blood cell pellet was resuspended in Cell Lysis Buffer. Once the pellet was completely resuspended and free of cell clumps, the solution was digested with RNase A for 15 minutes at 37°C. Proteins were precipitated by addition of the provided Protein
25 Precipitation Solution and removed by centrifugation. The DNA was precipitated out of the supernatant by addition of isopropanol. This method was simple and fast, requiring only 1-2 hours, and allowed for the processing of dozens of samples simultaneously. The yield of DNA was routinely >8 mg for a 20 ml sample of whole blood and had a MW of >50 kb. DNA was archived by storing coded 50 µg aliquots at -80°C as an ethanol precipitate.

30 DNA was genotyped using one fluorescently labeled oligonucleotide primer and one

unlabeled oligonucleotide primer. Labeled and unlabeled oligonucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, Iowa). All other reagents for microsatellite genotyping were purchased from Perkin Elmer-Applied Biosystems, Inc. ("PE-ABI") (Norwalk, Connecticut). Individual PCR reactions were performed for each marker, as described by PE-ABI using AmpliTaq™ DNA Polymerase. The reactions were added to 3.5 µl of loading buffer containing deionized formamide, blue dextran and TAMRA 350 size standards (PE-ABI). After heating at 95°C for 5 minutes to denature the DNA, the samples were loaded and electrophoresed as described in the operator's manual for the Model 377 DNA Sequencer (PE-ABI, Foster City, California). After gel electrophoresis, the data was analyzed using PE-ABI GENESCAN™ and GENOTYPER™ software. First, within the GENESCAN™ software, the lane tracking was manually optimized prior to the first step of analysis. After the gel lane data was extracted, the standard curve profiles of each lane were examined and verified for linearity and size calling. Lanes, which had problems with either of these parameters, were re-tracked and verified. Once all lanes were tracked and the size standards were correctly identified, the data were imported into GENOTYPER™ for allele identification. To expedite allele calling (binning), the program Linkage Designer from the Internet web-site of Dr. Guy Van Camp (<http://alt.www.uia.ac.be/u/dnalab/ld.html>) was used. This program greatly facilitates the importing of data generated by GENOTYPER™ into the pedigree drawing program Cyrillic (Version 2.0, Cherwell Scientific Publishing Limited, Oxford, Great Britain) and subsequent linkage analysis using the program LINKAGE (Lathrop *et al.*, 1985 *Am. J. Hum. Genet.* 37: 482-98).

5. Linkage Analysis

Fig. 1 demonstrates the pedigree of the individuals used in the genetic linkage studies for this invention. Specifically, two-point linkage analysis was performed using the MLINK and LINKMAP components of the program LINKAGE (Lathrop *et al.*, 1985 *Am. J. Hum. Genet.*, 37: 482-98). Pedigree/marker data was exported from Cyrillic as a pre-file into the Makeped program and converted into a suitable ped-file for linkage analysis.

The original linkage analysis was performed using three models: (i) an autosomal dominant, fully penetrant model, (ii) an autosomal dominant model with reduced penetrance,

and (iii) a quantitative trait model. The HBM locus was mapped to chromosome 11q12-13 by analyzing DNA for linked markers from 22 members of a large, extended kindred. A highly automated technology was used with a panel of 345 fluorescent markers which spanned the 22 autosomes at a spacing interval ranging from 6-22 cM. Only markers from this region of chromosome 11 showed evidence of linkage (LOD score ~3.0). The highest LOD score (5.74) obtained by two-point and multipoint analysis was D11S987 (map position 55 in Fig. 2). The 95% confidence interval placed the HBM locus between markers D11S905 and D11S937 (map position 41-71 in Fig. 2). Haplotype analysis also places the *Zmax1* (*LRP5*) gene in this same region. Further descriptions of the markers D11S987, D11S905, and D11S937 can be found in Gyapay *et al.*, 1994 *Nature Genetics*, Vol. 7.

In this invention, the inventors report the narrowing of the HBM interval to the region between markers D11S987 and GTC_HBM_Marker_5. These two markers lie between the delimiting markers from the original analysis (D11S905 and D11S937) and are approximately 3 cM from one another. The narrowing of the interval was accomplished using genotypic data from the markers D11S4191, D11S1883, D11S1785, D11S4113, D11S4136, D11S4139, (Dib *et al.*, 1996 *Nature* 380: 152-4), FGF3 (Polymeropolous *et al.*, 1990 *Nucl. Acid Res.*, 18: 7468) (information about the genetic markers can be found at the Internet site of the Genome Database, <http://gdbwww.gdb.org/>), as well as the markers GTC_HBM_Marker_1, GTC_HBM_Marker_2, GTC_HBM_Marker_3, GTC_HBM_Marker_4, GTC_HBM_Marker_5, GTC_HBM_Marker_6, and GTC_HBM_Marker_7.

As shown in Fig. 1, haplotype analysis with the above genetic markers identifies recombination events (crossovers) in individuals 9019 and 9020 that significantly refine the interval of chromosome 11 to which the *Zmax1* (*LRP5*) gene is localized. Individual 9019 is an HBM-affected individual that inherits a portion of chromosome 11 from the maternal chromosome with the *HBM* gene, and a portion from the chromosome 11 homologue. The portion inherited from the *HBM* gene-carrying chromosome includes markers D11S935, D11S1313, GTC_HBM_Marker_4, D11S987, D11S1296, GTC_HBM_Marker_6, GTC_HBM_Marker_2, D11S970, GTC_HBM_Marker_3, D11S4113, GTC_HBM_Marker_1, GTC_HBM_Marker_7 and GTC_HBM_Marker_5. The portion

from D11S4136 and continuing in the telomeric direction is derived from the non-HBM chromosome. This data places the *Zmax1* (*LRP5*) gene in a location centromeric to the marker GTC_HBM_Marker_5. Individual 9020 is an unaffected individual who also exhibits a critical recombination event. This individual inherits a recombinant paternal chromosome 11 that includes markers D11S935, D11S1313, GTC_HBM_Marker_4, D11S987, D11S1296 and GTC_HBM_Marker_6 from her father's (individual 0115) chromosome 11 homologue that carries the *HBM* gene, and markers GTC_HBM_Marker_2, D11S970, GTC_HBM_Marker_3, GTC_HBM_Marker_1, GTC_HBM_Marker_7, GTC_HBM_Marker_5, D11S4136, D11S4139, D11S1314, and D11S937 from her father's chromosome 11 that does not carry the *HBM* gene. Marker D11S4113 is uninformative due to its homozygous nature in individual 0115. This recombination event places the centromeric boundary of the HBM region between markers D11S1296 and D11S987.

Two-point linkage analysis was also used to confirm the location of the *Zmax1* (*LRP5*) gene on chromosome 11. The linkage results for two point linkage analysis under a model of full penetrance are presented in Table 5 below. This table lists the genetic markers in the first column and the recombination fractions across the top of the table. Each cell of the column shows the LOD score for an individual marker tested for linkage to the *Zmax1* (*LRP5*) gene at the recombination fraction shown in the first row. For example, the peak LOD score of 7.66 occurs at marker D11S970, which is within the interval defined by haplotype analysis.

TABLE 5

Marker	0.0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
D11S935	- infinity	0.39	0.49	0.47	0.41	0.33	0.25	0.17	0.10
D11S1313	- infinity	2.64	2.86	2.80	2.59	2.30	1.93	1.49	1.00
D11S987	- infinity	5.49	5.18	4.70	4.13	3.49	2.79	2.03	1.26
D11S4113	4.35	3.99	3.62	3.24	2.83	2.40	1.94	1.46	0.97
D11S1337	2.29	2.06	1.81	1.55	1.27	0.99	0.70	0.42	0.18
D11S970	7.66	6.99	6.29	5.56	4.79	3.99	3.15	2.30	1.44
D11S4136	6.34	5.79	5.22	4.61	3.98	3.30	2.59	1.85	1.11
D11S4139	6.80	6.28	5.73	5.13	4.50	3.84	3.13	2.38	1.59
FGF3	0.59	3.23	3.15	2.91	2.61	2.25	1.84	1.40	0.92
D11S1314	6.96	6.49	5.94	5.34	4.69	4.01	3.27	2.49	1.67
D11S937	-infinity	4.98	4.86	4.52	4.06	3.51	2.88	2.20	1.47

A single nucleotide polymorphism (SNP) further defines the HBM region. This SNP is termed SNP_Contig033-6 and is located 25 kb centromeric to the genetic marker GTC_HBM_Marker_5. This SNP is telomeric to the genetic marker GTC_HBM_Marker_7. SNP_Contig033-6 is present in HBM-affected individual 0113. However, the HBM-affected individual 9019, who is the son of 0113, does not carry this SNP. Therefore, this indicates that the crossover is centromeric to this SNP. The primer sequence for the genetic markers GTC_HBM_Marker_5 and GTC_HBM_Marker_7 is shown in Table 6 below.

TABLE 6

Marker	Primer (Forward)	Primer (Reverse)
GTC_HBM_Marker_5	TTTTGGGTACACAATTCAGTCG (SEQ ID NO:63)	AAAACTGTGGGTGCTTCTGG (SEQ ID NO:65)
GTC_HBM_Marker_7	GTGATTGAGCCAATCCTGAGA (SEQ ID NO:64)	TGAGCCAAATAAACCCCTTCT (SEQ ID NO:66)

The kindred described have several features of great interest, notably that their bones, while very dense, have an absolutely normal shape. The outer dimensions of the skeletons of

the HBM-affected individuals are normal, and, while medullary cavities are present, there is no interference with hematopoiesis. The HBM-affected members seem to be resistant to fracture, and there are no neurologic symptoms, and no symptoms of impairment of any organ or system function in the members examined. HBM-affected members of the kindred live to advanced age without undue illness or disability. Furthermore, the HBM phenotype matches no other bone disorders such as osteoporosis, osteoporosis pseudoglioma, Engelmann's disease, Ribbing's disease, hyperphosphatasemia, Van Buchem's disease, melorheostosis, osteopetrosis, pycnodysostosis, sclerostenosis, osteopoikilosis, acromegaly, Paget's disease, fibrous dysplasia, tubular stenosis, osteogenesis imperfecta, hypoparathyroidism, pseudohypoparathyroidism, pseudopseudohypoparathyroidism, primary and secondary hyperparathyroidism and associated syndromes, hypercalciuria, medullary carcinoma of the thyroid gland, osteomalacia and other diseases. Clearly, the HBM locus in this family has a very powerful and substantial role in regulating bone density, and its identification is an important step in understanding the pathway(s) that regulate bone density and the pathogenesis of diseases such as osteoporosis.

In addition, older individuals carrying the *HBM* gene, and therefore expression of the HBM protein, do not show loss of bone mass characteristic of normal individuals. In other words, the *HBM* gene is a suppressor of osteoporosis. In essence, individuals carrying the *HBM* gene are dosed with the HBM protein, and, as a result, do not develop osteoporosis. This *in vivo* observation is strong evidence that treatment of normal individuals with the *HBM* gene or protein, or a fragment thereof, will ameliorate osteoporosis.

6. Physical Mapping

To provide reagents for the cloning and characterization of the HBM locus, the genetic mapping data described above were used to construct a physical map of the region containing *Zmax1* (*LRP5*) on chromosome 11q13.3. The physical map consists of an ordered set of molecular landmarks, and a set of BAC clones that contain the *Zmax1* (*LRP5*) gene region from chromosome 11q13.3.

Various publicly available mapping resources were utilized to identify existing STS markers (Olson *et al.*, 1989 *Science* 245:1434-5) in the HBM region. Resources included the

GDB, the Whitehead Institute Genome Center, dbSTS and dbEST (NCBI), 11db, the University of Texas Southwestern GESTEC, the Stanford Human Genome Center, and several literature references (Courseaux *et al.*, 1997 *Genomics* 40: 13-23; Courseaux *et al.*, 1996 *Genomics* 37: 354-65; Guru *et al.*, 1997 *Genomics* 42: 436-45; Hosoda *et al.*, 1997
5 *Genes Cells* 2: 345-57; James *et al.*, 1994 *Nat. Genet.* 8: 70-76; Kitamura *et al.*, 1997 *DNA Research*, 4: 281-9; Lemmens *et al.*, 1997 *Genomics* 44: 94-100; and Smith *et al.*, 1997 *Genome Res.* 7: 835-42). Maps were integrated manually to identify markers mapping to the region containing Zmax1 (LRP5).

Primers for existing STSs were obtained from the GDB or literature references are
10 listed in Table 6 below. Thus, Table 7 shows the STS markers used to prepare the physical map of the *Zmax1* (*LRP5*) gene region.

TABLE 7: HBM STS Table

STS Name	Locus Name	Type	Gene Access. #	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
ACTN3		Gene	DB:197568	0.164	67: CTGAGCTAGTGGCTCTCT	68: TTGAGACACTTGGCTGG	Actinin, alpha 3 - skeletal muscle
PC-8/PC-Y		Gene	DB:197884	0.125	69: CTGAGTGCATGAGATGGA	70: CAGATGATGATGCTCAGG	Pyruvate Carboxylase
ADRBK1	D11S2161E	Gene	DB:4590179	0.322	71: GTTTCAGGAGCTCAGATGTC	72: TTCTGGAGGTTGCTGTGAG	Adenosine Receptor (A2) Gene
PSANK3		Gene	DB:197566	0.117	73: TTATGTGATTTCCGTTGGC	74: GCTCTGTCTGCTGACTCAGG	Beta-adrenergic receptor kinase
PP1(12)/PP1(22)		Gene	DB:197566	0.259	75: GAGAAAGAAATTAAGGGAGC	76: TCGTTTGAAGCACTGAGGA	sim. to Human endogenous retrovirus mRNA long terminal repeat
GSTP1-PC81		Gene	DB:270066	0.208	77: GAGATGCGGAGTTCATGTGGCT	78: ATACACGAGCTCATGTTTCCCGT	Protein phosphatase 1, catalytic subunit, alpha isoform
NDUFV1		Gene		0.139	79: AGGCTGGGCCACAGGTGAGTACGT	80: TCCGAGGACTTGACACCGCTTCA	Glutathione S-transferase pI
PSANK2		Gene		0.521	81: CATGTGCTCCCACTCATCTAT	82: CAGATGCTGTAGCTCTTGG	NADH dehydrogenase (ubiquinone) flavoprotein 1 (51kd)
PSANK1		Gene		0.157	83: CAGAGAGTCAAGGAGCTTG	84: ATCTCTCATCATCCATCTTT	Aldehyde Dehydrogenase 8 (ALDH8)
UTS20	D11S1917	MSAT	DB:314521	0.211	87: AAGTCAGGCTGCAAGAG	88: TAGGAGCACTTCACTCTTT	Human ribosomal protein L37 (PSANK1) pseudogene
AFM289a9	D11S1337	MSAT	DB:199805	0.287	89: AAGTGTGAGGATCTCATG	90: ACTCATGAGGGCTATTAT	
GALN		Gene		0.322	91: GCTTCTGCGAGTGTATCAAC	92: ATGCGAGAGAGCTATAGACA	Preprogalanin (GAL1)
BM551		VNTR	DB:177850	93: GATCAGCGAATCTCTCTCGGCTC	94: TCCATGATGAGGACTTGGGAAG		
BCL1(1)/BCL1(2)		Gene		0.205	95: GCTATTCAGACTTACGCA	96: TTGCATCTGCTTGGATGCA	B-cell CLL/lymphoma 1 - Cyclin D1 (PRAD1) gene
CDND1		Gene	DB:4590141	0.248	97: GCAAGCTGTAGTGGGTTCTAGGC	98: CAGCGCAAGAGGACATGCAACGGC	Cyclin D1
FGF4		Gene	DB:4590113	0.549	99: CACGATGAGTGGGTTCAAGGAG	100: CAGCAGAGAGTCTCCAGGCTATAC	Fibroblast growth factor 4
FGF3-PC81		Gene	DB:188627	0.161	101: TTCTGGGTTGTCTGAAT	102: ACACAGTTGCTCTTAAAGGCT	Fibroblast growth factor 3
AFM1642F12	D11S913	MSAT	DB:188151	0.22	103: CATTTGGGAAATCCAGAGA	104: TAGGTGCTCTATTTTGTGCTTC	
AFM190Y05		MSAT	DB:1222329	0.275	105: GACATACATGACATATAAGAGG	106: CAACCATACAGGAGTAAG	
D11S4689		STS	DB:740600	0.147	107: GACAGAGGGGTAGTTGGC	108: TGGGACACAGATGATGATGG	
D11S4540		STS	DB:740102	0.167	109: GAGTGTCTCTCTTAAATCTTTG	110: GAATCTATTGTAGTATGAGGAG	
D11S4664		STS	DB:740516	0.158	111: CTGTGACCCCAAGTCC	112: TCTGCTCTTAACTTCTCGG	
D11S4327		Gene	DB:474522	0.311	113: ACTCATCATCTCATCACTG	114: TGTGCTTGTCTCTCTGAC	Choline Kinase
D11S703		STS	DB:196290	0.166	115: GTGGAGGAGTATGAGG	116: TGTGCTCTCTGCTCTGAG	
AFM147XD10	D11S1889	MSAT	DB:307895	0.183	117: AGCTGAGCTCTACAGATG	118: CAGGAGCTGTGAGAGGCTG	
AFM1431Y05	D11S987	MSAT	DB:195002	0.089	119: GACTCGAGTGTGGCAATAAAGC	120: GTGGCAGCATGACTCTAAAG	
AFM0358a9	D11S4178	MSAT	DB:611922	0.237	121: CAGGCCAGCTCTCTTG	122: CGTGTCACATGAAGTGG	
AFM272a95	D11S4113	MSAT	DB:608115	0.218	123: ACTCACAGGTGTAATCC	124: CTTGAGCCCATCTTGG	
WT-17803		EST	DB:4581644	0.15	125: TATTTGCAAGCTTGAAGTCTCT	126: ATCTGCTGCTGTGTGTGCC	
SGC31923		EST	DB:4578606	0.286	127: ACTTTTGTCTGAGGTGGC	128: ATCTGCTGCTGTGTGTGCC	Transformation-sensitive protein 1ER SSP 3521
D11S4364		Gene	DB:4578606	0.324	129: GAGCAGGAGGAGAGGC	130: CCAAGCTGCTGTGTATTTG	ZNF162 - splicing factor 1
SGC3523		EST	DB:4582598	0.13	131: AGCACTTATTTGTTTATTTGATC	132: AAGAGTGAACAAAGCAACATCC	
WT-16754		EST	DB:4578377	0.15	133: GTGGAGTGTGGGATGAGG	134: TACTGTTCTGATGATGATGCGG	
WT-6315	D11S4418	EST	DB:4578804	0.224	135: ATGAGTGTGGATGATCTAATATT	136: TCCCTCAAGAGATGTAAGG	
WT-16915		EST	DB:4584055	0.125	137: CTGGTCTCTGTGTGCTG	138: ATCCACCAAGGCGAGGAT	
SGC30608		EST	DB:4584055	0.128	139: TCAAGAGAGAGTGTATTAACA	140: CTGCTTGAAGTGTATGAGGCC	Mitogen inducible gene (MIG-2)
WT-17663		EST	DB:4583346	0.126	141: CAGCCGAGTGTATTTATGAAA	142: GATCCAGGAGCATGGAC	
SGC31567		Gene	DB:1222327	0.199	143: GATATAGAACATTTATTTGCG	144: CTGAGAGGAGGAGCAGAG	Human lat interactive protein (TIF60)
SGC30658		Gene	DB:4578432	0.207	145: CTACACACACACAGGC	146: CAGGAGAGCTGCTCTTC	Calcium activated neutral protease large subunit, nuCANP, calpain
SGC34590		EST	DB:4584037	0.13	147: GTGCTTCTGACTTCAAGTCTGTC	148: TTTCTCTCAACATCACTACTCC	
SGC33927		EST	DB:4582382	0.15	151: TAAATATCCAGCTTAAAGCAT	152: AGCTTCCAGATGAGGCC	
WT-18671		EST	DB:1222325	0.124	153: TGGTTTAAACCTTTAATGAGAAA	154: TGTGATCTATACCTGTTTCCG	
WT-12334		EST	DB:1222257	0.127	155: AATATTTAAAGAGAGGAGAGGCA	156: TGGCTGTGAGCTCTCTGGA	
WT-18402		EST	DB:4581874	0.113	157: GTTACAGAAACATTTTGAAGAT	158: TGAAGCTTATGTTCTCTCTG	
WT-18671		EST	DB:4584947	0.131	159: TCGAAACCATTTATTCACCG	160: TCTGGGCTGTGTGATTT	HLK
WT-12856		EST	DB:4576606	0.209	161: TCGAAACCATTTATTCACCG	162: TGTCTCTCTCCAGCAGG	HLK
SGC33767		EST	DB:4581106	0.15	163: CTTTATTAACATTTGATGTCGA	164: TGTCAATTTCCGCCAAA	
AFM343Y85		MSAT	DB:1222332	0.181	165: AAACACAGCCACCA	166: CCTTGGAAAGGTAAAGTCT	
SGC33744		EST	DB:4578826	0.15	167: CTTTTGGTAGACACAGGCTCA	168: TATCTGTGATGCTTCAATGT	
SGC32272		EST	DB:4581592	0.135	169: GAGCAAGGTGATTCAGGCG	170: ACTGAGAGTCTGTCTCT	
SGC34148		EST	DB:4580084	0.11	171: CAGTAAAGAGTCACTATGGCTCA	172: CACTTCTCCACTTGTGCC	
WT-18546		EST	DB:4574596	0.133	173: TTATGATGAGTATGTAAGCC	174: TGGCAAGTATGAGCAGATCA	Human 1.1 kb mRNA upregulated in retinoic acid treated HL-60 neutrophilic cells
SGC31103		EST	DB:4582765	0.175	175: CTATCCAGAGATGAACAGG	176: TCCATTAAGGCTATGCTCC	
SGC30028		Gene	DB:4580505	0.128	177: GCACATTTATGATTAACCTCC	178: CACTGGAGTACAGTGGTGG	Human pyruvate carboxylase precursor
WT-2875	D11S4407	STS	DB:4578546	0.125	179: CATCCACATCACTCATG	180: GGGGAGTCTTAAGATTTTGA	LAR-interacting protein 1b
SGC36905		Gene	DB:4577182	0.233	181: AGATACATTTTGGACCATGG	182: TGAAGGATATTTAGCTCTGGA	
GCT16607	D11S4270	STS	DB:626245	0.137	183: GAGGTTTGTCTCTCTGATC	184: TGAAGTGTGGGAGATCA	
WT-6504	D11S3974	EST	DB:458142	0.174	185: CTCTCATGACCAACCCG	186: CACTACTGTTGACATGCCA	
SGC31049		EST	DB:4580093	0.15	187: TCTTACTGCTTACAACTTCT	188: CACTGAGCTGCTGATCTG	
TIGR-A00217		EST	DB:1222183	0.199	189: AGATCAGCAGGAGATG	190: CATTCACATGATGATG	NDUFV1
WT-5996	D11S2282	EST	DB:458683	0.11	191: CATACCTATGAGTGTGCTACAGG	192: GATTTTCTCATCTCTG	amphipin (ENS1)
WT-16897		EST	DB:4578848	0.15	193: TTACAGCCACAGGTTTCC	194: AGGTGTGTGCTGCTGGA	Nuclear mitotic apparatus protein 1, NUMA
SGC31912		EST	DB:4567868	0.101	195: CACTGTATCTCACTTAACCTGAGG	196: TTGATTTGTTGCTGCCAAA	

TABLE 7: HBM STS Table

STS Name	Locus Name	Type	Gene Access. #	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
WT-13500		EST	GD8:4577893	0.15	197: CCCACCTCCACTTTTATTT	198: CCAGTCACTTTACTAGTCTTTTG	
CHLC-GAAT1801.P7933	D115971	MSAT	GD8:684255	0.103	199: AGGACACAGCTGCTCATCTAG	200: ACCAGCATTTGCACTAAAG	LAR-interacting protein 1a mRNA
SGC35519		Gene	GD8:4577180	0.134	201: GATGGTCACTAATCTGTCA	202: ACATTTATTTGCACTCAAC	Carboline palmitoyl transferase 1
WT-11974		EST	GD8:1222255	0.108	203: AGCATCTTATGCTGACGCA	204: ATGCTGCTGGCTGGAAG	Beta-adrenergic receptor kinase 1, ADRB1
WT-15244		Gene	GD8:4574740	0.108	205: TACATTCAAATTCGGGAA	206: CTGCTGTGTGTGTGCG	
WT-17496		EST	GD8:4583336	0.131	207: TGTTTATTTCTCAGTAAAGGCA	208: GACTCTCTGTGACACACG	FGF4
WT-9159	D1154381	EST	GD8:678144	0.111	209: CACCAATATTTATATGTTCTGCG	210: GTAAGATCTCTCACTGTGTGACC	
WT-4232		STS	GD8:1222250	0.175	211: CCTATATGGCTGACGAA	212: ACTCTCATGTGAGTCAACG	Human DNA helicase gen (SMBP2)
SHGC-4167		EST	GD8:4566789	0.161	213: CAGTGTGCAAGTTTCTATTT	214: CAGCATCTTCAAGCTTACC	
WT-14303		EST	GD8:4576938	0.15	215: CTGCAATTTATGAGATTAACAG	216: TGTCTGCTGGAGTCTAGAGTTC	
WT-16597		EST	GD8:4585666	0.13	217: CAGGCACTGAGACACTTACC	218: AAGGATCAAGGAGGCACTTTC	
RC95CATTFOR/KC95CATHREV		MSAT	GD8:191084	0.15	219: ACACATCTCTCTGTCGCC	220: TGAACCTTGGAGGACAG	
WT-1282	D1151296	MSAT	GD8:198525	0.362	221: CATTCGCAAGTTTCTGAC	222: GTCTGGGATATACAGTGT	
D1151599E		EST	GD8:335216	0.07	223: CAGAGAACTGCTGTAGCC	224: CCATGCTAGAGCCCTGGGTTTC	
D115468		STS	GD8:4585666	0.096	225: AGTGTGGGAGCACTCTG	226: CAGACAGATAGCCCTGGGTTTC	
D115668		STS	GD8:179349	0.143	227: TCCCTCATCTCTCTCTCTG	228: AGCCCTCCCTGGGATATC	Aldehyde dehydrogenase (ALDH8)
RH18048		Gene	GD8:4572853	0.188	229: GATGCTTACTTACACGCG	230: AGGATCTCTATCTGGCTATG	Human DNA helicase gen (SMBP2)
IGHMBP2		Gene	GD8:4590087	0.699	231: TGCGACACATGCTCGGCTG	232: GAGAAGCCGGGAGGCTCTG	Nuclear ribotic apparatus protein 1, NUBA
NUMA		Gene	GD8:4590244	0.277	233: CTGATCAACACAGATTTGAGGCT	234: GGGTGTGAGCTGCTGCTGAAGG	High sulphur keratin, KRN
KRN1		Gene	GD8:4590232	0.228	235: AGTGGAACTCAGTACTCCGGA	236: CAGTTTGGCTCAGCATATGAGGGCA	
Cbl1006	D115202E	EST	GD8:445887	0.091	237: CATTAAGTAGTGGGGGACAG	238: CAAGAGCAAGTGAATAGGG	protein phosphatase 2A, PP2A
RH10753		Gene	GD8:4563588	0.194	239: GAGTAGACATGATTTACTG	240: CATGCTATTTATTTCTCG	Amplixin
EM51		Gene	GD8:459016	0.64	241: CGCCCTGGATCTCTCACTACA	242: GGGCATCAGGAGGATGGGTAGA	Androgen Receptor
INPPL1	DX59736	Gene	GD8:737674	0.137	243: GCTCTATCTGTGTGTGATAGG	244: ATGGCACTGACTTCTGCTCTG	51C protein, Inositol polyphosphate phosphatase-like 1
RH18051		Gene	GD8:4590093	0.382	245: CTGGAGGCTATGAGGAGGCG	246: CAGCACTACTCTTGGGG	NOF1
SHGC-11098		EST	GD8:4572859	0.195	247: TTGGAGTCAAGGCGG	248: GATGAGCAACTGCTGTGAC	
Cbl1011	D115297E	EST	GD8:445869	0.11	249: ACAAGCTGCTTAGACCTG	250: GATGAGCAACTGCTGTGAC	
D1151250		EST	GD8:335210	0.247	251: TTTTCAATATGTACTTC	252: CAATCCCAACCGTACAGGCG	
NDUFV1	D115245E	EST	GD8:445695	0.233	253: CTGATCTCGCCGAGGAC	254: GCTCTGCTGAGGATGAGAC	NDUFV1
AFM0322g5	D1154136	MSAT	GD8:609546	0.19	255: GATCGCTGAAACCCAG	256: CCAGTGGTCTTAAAGCG	
AFM059xg9	D1154196	MSAT	GD8:614025	0.2	257: GATCGCTTCTTCTGAGGCT	258: TAATGTGCTGCTGCG	
Cbl1712	D115288E	EST	GD8:445842	0.158	259: AGGGAATATGTTATGAGGAG	260: GAGTGTGTGAAGGAGG	
SHGC-1364		EST	GD8:4562765	0.137	261: ATGACACAAATGAGCAAAACAGG	262: CCAGCAAGTTTCTCTCAATCC	
RH17410		EST	GD8:4571587	0.126	263: TGACATCTTGCATATGCG	264: AGTATCCCACTGATACCG	
RH17414		EST	GD8:4571595	0.121	265: AGCTCTGCTCTGAGTCA	266: CAAGTGTCTCTGTTGTTG	
RH17770		EST	GD8:4572301	0.267	267: GCTCTCAAGTATGTTGAAACC	268: TGTGTATCATAGTAGTGAACAG	513 avian erythroblastosis oncogene homolog
SEA		EST	GD8:4590169	0.13	269: CTGAGGCCAGGCACTACT	270: GGAATCTTCACTGACGAGT	
RH10689		EST	GD8:4563460	0.107	271: AATGATGATCTCACTCTG	272: ACTGAGAACTCTTGTCT	
TGR-A006P20		EST	GD8:4587692	0.236	273: GACATCTGTATCTCATATTC	274: GGTACAGTGTCTGCTT	
TGR-A007D15		Gene	GD8:4586398	0.24	275: CTATGTCAAAACAGAGAGG	276: ATCTAGTTTCTCTCTCTT	Mein gene (MEIN1)
TGR-A008B14		EST	GD8:4588882	0.141	277: GTAAATGAGAAACAGCAAAATGA	278: CTATTGATGTGATATGTTATGG	
TGR-A008P15		EST	GD8:4589094	0.203	279: AAGTAAACAAATAGAGGGAC	280: CTTACCCCAAGGTAACAG	
TGR-A008T11		EST	GD8:4589662	0.182	281: ACTTCTATAATGAGGTGAG	282: GAGGAGCTTCAAGAGGAA	
TGR-A008U48		EST	GD8:4589278	0.138	283: CATCTCTAGACTCAAGGATC	284: GAATGATGATGATTTCTTTG	
TGR-A008K45		EST	GD8:4589364	0.107	285: GTGTGAGGAGAAAGCACT	286: CTCGAGTAGTCACTATCC	
SHGC-11839	D1154611	Gene	GD8:740339	0.242	287: CAAGTACAAATTAACCTTAAGCCG	288: CAGACCTCTCTCTCAAAAC	
NIH1242		EST	GD8:3889276	0.149	291: TTTCTAGTCAAAAGCGTTC	292: CCATGGCTTCTCTCTT	Folate receptor 2 (FBR2)
SHGC-11599	D115429E	Gene	GD8:737558	0.147	293: CACGACAGGTTGGGCTG	294: ACTATTAGACATGAGCGCG	cgMP-stimulated 3'-5'-cyclic nucleotide phosphodiesterase PDE2A3 (PDE2A)
SHGC-11867	D1154331	Gene	GD8:674684	0.14	295: CTATGCTGATGACCC	296: TGCCTTCTTGAACCTTATTC	Macrophage Migration Inhibitory factor
SHGC-15349	D1152124	EST	GD8:740819	0.141	297: TACAGCTTCACTGACGG	298: ACATGCTGTGGACATG	P2U Putrinceptor
B4B8a05	D115235E	EST	GD8:445662	0.095	299: CTGAGCTACTGCAAG	300: CCTGACTTGGACATGTC	
B689a07	D115238E	EST	GD8:445674	0.09	301: TGAAGTCACTCTGCTCC	302: CAATTAAGCTCTCCAGAC	
Idi1		Gene	GD8:197840	0.23	303: GGGATTTTCTCAGGAC	304: GGTGAGAGTGGGCAAT	Folate receptor-2 (FBR2)
NIH1738	D1154284	EST	GD8:626560	0.173	305: TTTCAATTTAGAGCACTG	306: CTTAAGCCTGTTTGTG	Folate receptor-3 (FBR3)
WT-7351	D1154433	Gene	GD8:679143	0.324	307: CTCTCACTCACTGCAAAAC	308: TGGAGAACCCCAAGAGAC	
WT-11325		EST	GD8:4578507	0.12	309: AAGACCAAAATGACACAA	310: GTGTGTGGGCAACAATTG	
WT-15192		EST	GD8:4579806	0.19	311: AGAGCACTTCTCTGAC	312: AGAATCTCATCAGGGCG	
WT-17872		EST	GD8:4577492	0.141	313: AAAAAAGAGGTCTTAAATTTGA	314: AATGTTTTTTTGTGTTTATGAGT	
SHGC-30732		EST	GD8:4567830	0.105	315: GATTTAGGAGATGACAAATGCGG	316: GGGGCAAAATTAATCTTTATTCAGG	
SHGC-288		EST	GD8:4566057	0.123	317: CATCATCATTTGTTGAC	318: TGCTCTCCCAAGAGAG	
WT-13814		EST	GD8:4579290	0.15	319: TTAAGTGCATTAACCTATGAC	320: CAGAGAGATGACCAATGCG	DRES9
WT-14122		Gene	GD8:4576114	0.128	321: CCATCTCTTTATGAGGTTGG	322: CTCTGTGCAAGTAGCATCTTACA	Human VEGF related factor isoform VRF186 precursor (VRF)
2729/2730	D1154057	EST	GD8:596509	0.118	323: CCACTGTGTTATTTCCAGC	324: AGAGCCCATATCAATGAC	
SHGC-31329		EST	GD8:4567286	0.15	325: AGCTTAAAGTAGCAACCACTGG	326: GGAATCTTCACTCAAGAG	

TABLE 7: HBM STS Table

STS Name	Locus Name	Type	GDB Access. #	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
SGC3858		EST	GDB:4578600	0.127	377: TGTGTTTATTTCACCTGCC	328: AGAGTGGCTGCAGCCGAG	
WI-12191		EST	GDB:1222208	0.15	329: TTTTCTTTTATACAGAAATTTGAGG	330: TGAGGAAGTAAACACAGCTCATC	
WI-13701		EST	GDB:4574892	0.15	331: ATGAATCTTAAGCAAGATCCCA	332: CACAGAGTCCAGGGTCTGT	
WI-14069		EST	GDB:4584373	0.15	333: AAGGCTTTTATTATCTCTCTG	334: GCTCTAGAGCTGCTGGGT	
WI-14272		EST	GDB:4578523	0.125	335: GCTCTTAAGTCTTAGAGTCAGCTGG	336: AGCCCAAGTACAGCTATCC	
WI-17347		EST	GDB:4578523	0.127	337: TTGGTAAATGATGCCACAGA	338: TGGTCCCACTCACTATCC	
WSG1561		EST	GDB:4564415	0.215	339: ACACAGCATGCGAGGAGAG	340: ATCCCTGGTCTTAGGTGG	
WSG1938		EST	GDB:4564568	0.137	341: GATGGAAGTAGCTCTCTCTGG	342: GGAAGGACAGCAAGTACTACC	
WSG2259		EST	GDB:4565137	0.141	343: CCGGTGCTTGGGAAGATG	344: GAAGTCTCTGTTGGGGGA	
RP97		EST	GDB:4559690	0.17	345: TTACAGCATAGTCACTATAGC	346: ACACCTCTACAGCCCTTACA	
WSG4794		EST	GDB:4573113	0.141	347: CCTCTCTCACTCACTAC	348: GCTCACTTCTTGAAGGC	
WSG4957		EST	GDB:4569051	0.171	349: AGATAGGCAACACACTGG	350: GTTGAATATAGACAGGGGCC	
WSG4974		EST	GDB:4569051	0.166	351: TTCTAGGTCAAGGCTGTCT	352: AGTTGGAAATCTGCTGTCA	
WSG8144		EST	GDB:4573137	0.17	353: ACTGAGTCCCTCCACCC	354: TCTCTCACTCTCTCCAGA	
WSG9175		EST	GDB:4569999	0.19	355: GTGATCAGCGGTCACTCTG	356: TGGAGAGCTGCTGTGACCC	Human protein kinase (MLK-3)
SHGC-10667	Gene	EST	GDB:740246	0.277	357: CTGACAGTGGCTCAAGTTTC	358: TCAAAAGTCTGGTGTGACGC	FGF3
SHGC-11930	Gene	EST	GDB:1231223	0.21	359: ATTTCAGAGCTAGCTCAA	360: CTTTATGTGTGTGATGACACAAAGC	
SHGC-32786	EST	EST	GDB:4567878	0.125	361: GATCATGCACTGTGACAC	362: TACATTGMACTTTAAMACTGA	PK916-Binding Protein Precursor (FKBP-13)
FRBP2	Gene	EST	GDB:4585059	0.064	363: AACTGAGCTGTAACTCAGACTGGGA	364: TGGACAGTCTGCTCTGATGG	
MDJ1	Gene	EST	GDB:4585064	0.202	365: TTATCCCTTATTTCTCTCTTG	366: TGGTCACTGATTTATTCTGAGG	
S453	Gene	EST	GDB:4585064	0.859	367: TCTTCAAGCTCTGCTAGTACC	368: CTCACTCAACTGCTTAAACC	4F2 Cell-Surface Antigen Heavy Chain (4F2HC)
STS1-CSRL-112e11	STS	STS	GDB:196276	0.108	369: GTGGCTCAGCTATGTAGACAC	370: CAGCAGACATGCGCTAAGTCC	
STS1-CSRL-44a3	STS	STS	GDB:547681	0.135	371: CTGATTGAGACACAGACAG	372: TAAAGCCCTATACCTCTCC	
STS1-CSRL-310b2	STS	STS	GTC:547609	0.118	373: TAGTAGGAGCTTCAACAG	374: AGATGTTTGGTATGACTTGG	
STS1-CSRL-409	STS	STS	GDB:459728	0.123	375: GATGATTAACTCTCTGGC	376: GAGACAGTAAGCACTCATG	
STS1-CSRL-10323	Gene	EST	GDB:197824	0.196	377: GAGTGTGGGACCTGTA	378: AGAGGAGGAGAACACACTT	Folate receptor 2 (FRP2)
WI-9219	Gene	EST	GDB:676135	0.141	379: GACAGATCTGCCAGAG	380: TCCCAAGCTCTATCCAC	Collagen binding protein 2, collagen-2 gene (CBP2)
GTC_ZNF	Gene	EST	GDB:678179	0.1	381: GGAGGATGAGCAAGTCTGA	382: GTCCAGCTGTGACTATCC	Retinal outer segment membrane protein 1, ROM1
AFMa152yH1	Gene	EST	GDB:603797	0.172	383: TCAAAACAGTCTCTCCA	384: GCAAGGCTTTACCATATTC	ZNF126
AFMa1312h5	MSAT	EST	GDB:611241	0.263	387: GTTCTCAGACAGACAGAC	388: GAGACAGCACTATTGCCC	
AFH03Bv69	MSAT	EST	GDB:609621	0.151	389: TATAGACTCAGCCCTGCTGC	390: CTTCTAGAGTCAAGTGTGG	
AFM12Dx63	MSAT	EST	GDB:199292	0.209	391: TTCTAGGACTCTCTTACT	392: CTCCCTCTGCTCAGTTTAT	Serine/threonine kinase
WI-18813	EST	EST	GDB:199292	0.13	393: ATCTAGCAGAGAGAGGCC	394: CTCCCTCTGCTCAGTTTAT	
WI-19549	EST	EST	GDB:199292	0.252	395: AACTTCTATTTCTCCAAAGGA	396: AGCAGATCTCTCTGCGCAT	
WI-20154	EST	EST	GDB:4583084	0.25	397: AACAGTGTGATCGGTAGGCA	398: AAGATATGAATGGATGGAGC	
WI-22393	EST	EST	GDB:4583084	0.142	399: GTGAGGTGGGCTTATTT	400: CCTATATCTGCTGCTGCC	DRES9
WI-7587	EST	EST	GDB:1223732	0.274	401: GCTCTAGTGGAACTGAGG	402: GATCTCAAGGCTCTGCTCTG	Ultra high-sulphur keratin protein (GRN1)
EST455579	EST	EST	GDB:1223732	0.273	403: GTTCTAGTCTCAAGGCAAA	404: CCAGTATGATGTTGTTCAACA	
WI-21134	EST	EST	GDB:1223732	0.293	405: GCTGCTTGGAAATTTCTGTT	406: GCTGCTGCTGGGGAAG	Fas-associating death domain-containing protein, FADD
WI-21698	EST	EST	GDB:740182	0.25	407: ATTCAGCTCATCAGACCC	408: GGACTGCGCTTTTGAAGTCT	
SHGC-37373	STS	STS	GDB:740182	0.225	409: ATATGACCTGTCACAAATAGS	410: AGACTGGGAAAGTGGAGAA	
SHGC-36533	STS	STS	GDB:740182	0.125	411: ATTGCACTGGAAATGCTT	412: TTAATCTTTTGTCACTCTCTGATT	Arix homeodomain protein, neuroendocrine specific, tx factor
ARX	Gene	EST	GDB:6262613	0.242	413: TGGTCTCTTCTTCAAGGAGC	414: GGAAGAGAGAGCTGCTGCTGAGGA	Chloride channel current inducer, [CLN] gene
CLIC1R	Gene	EST	GDB:6262613	0.242	415: TCAAGGCTGTGTTGCTGCTGCTG	416: AGCAGTAAAGGTTACAGTGGCGAGG	
B18N21-HL	STS	STS	GDB:6262613	0.242	417: AGCAGTCAAGCTCTCTTA	418: CCGGAGGAGAGATCTAT	
B23H17-HR	STS	STS	GDB:6262613	0.242	419: TGGTAAAGCAGAAATGTC	420: AATGGATGGGGGATATT	
B23G10-HR	STS	STS	GDB:6262613	0.242	421: CTGAGCTGTATCTGCTC	422: AGAGGCCAGTCTCAAGAT	
B24F23-HR	STS	STS	GDB:6262613	0.242	423: ATCACTGAACTGCACT	424: CCTCTCTTTTCTCTGTTT	
B337H24-HL	STS	STS	GDB:6262613	0.242	425: CAGGCTTTGAGAGAGAG	426: TAGAGCTTTAGTGAAGAC	
B337L5-HL	STS	STS	GDB:6262613	0.242	427: GCTCTGCTGCTGGTAAAA	428: GCTCTCAAGACTGTGCG	
B382N10-HR	STS	STS	GDB:6262613	0.242	429: CTTCTGCTGCTGGGAGAT	430: GACACTTGGGAGAAC	
B12J1-HR	STS	STS	GDB:6262613	0.242	431: GCTATGATCCCATCTG	432: GATCAGTCAATGAAGG	
B18D17-HR	STS	STS	GDB:6262613	0.242	433: TGGATACAGGGGTCAC	434: CCGAGACTGAAGATGA	
B23G66-HR	STS	STS	GDB:6262613	0.242	435: ACTGCTCTCTCTCTG	436: TGTCTTCTCTCTGGA	
B27B22-HR	STS	STS	GDB:6262613	0.242	437: ATGACAGCAAGCAATGT	438: TGTCTTCTCTCTGGA	
B31F21-HR	STS	STS	GDB:6262613	0.242	439: GCAAGGCTCTCTGAT	440: TTTGAGGATTCATCTT	
B337H24-HR	STS	STS	GDB:6262613	0.242	441: GCACATCTCTCTGAGG	442: ACTTTGCACTGTTGTTT	
B35B19-HR	STS	STS	GDB:6262613	0.242	443: GCACTTCTCTCTCTG	444: TGTCTTCTCTCTG	
B14N18-HL	STS	STS	GDB:6262613	0.242	445: AAGCTCTCAGAGAGAGA	446: GAGTCACTTGAACAGA	
B172N12-HL	STS	STS	GDB:6262613	0.242	447: AGGCATCAAGCTTCTCT	448: GGTTTAGAGACCGAGCC	
B172N12-HR	STS	STS	GDB:6262613	0.242	449: GTGCTGCTCAAGTATCC	450: GGAATCCCTTCTTCTCA	
B215111-HR	STS	STS	GDB:6262613	0.242	451: GACATTTGTTAGCAGC	452: GATGCTGTGAATGAACAA	
B22F55-HR	STS	STS	GDB:6262613	0.242	453: CTCAGCTCTGTTCTATGC	454: GCTGTGAGTGTCTGCTCT	
B312B3-HR	STS	STS	GDB:6262613	0.242	455: TACAGAAACCGCAGTCT	456: GCAACAAACCGCAGTCT	

TABLE 7: HBM STS Table

STS Name	Locus Name	Type	GBS Access. #	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
B328G19-HL		STS			457: AAAGGAGGAGATCATGG	458: TCATCTAGCAGGAGCAG	
B328G19-HL		STS			459: CTGAGCATCGATGATGAGC	460: GTGCAAAATGAGCAGCTT	
B32910-HL		STS			461: TCTAACCTTACTGGGC	462: TCTAACCTTGGGAATGA	
B32910-HL		STS			463: TTACACAGCAGCAGGGA	464: ATCTCCCTCCCTCAGAG	
B368G19-HL		STS			465: GTCCAGGGCTTATCT	466: TGAGCATAAATTTTCTAGCTG	
B368G19-HL		STS			467: GGAAGAGCAAAATATCCA	468: GTTGACAGANTTGTTCAT	
B368G19-HL		STS			469: AGCAGCTTATTCATGG	470: GTACACACAGCAGGAGCA	
B368G19-HL		STS			471: TCTGTGATATGAT	472: GGGGCTGAGAGTAGGAA	
B368G19-HL		STS			473: ATGGGATTAATACGG	474: AGTAGCATTGGCTCTT	
B368G19-HL		STS			475: CTGAGGAGAGAGCTGG	476: CGCTTACAGGCAAGTA	
B368G19-HL		STS			477: AGATGCTTCTAGGGT	478: CACATGCTCTCTCACTAA	
B368G19-HL		STS			479: GTCTAGAGCCCTTAA	480: CAATGCTCTCTCTCACTAA	
B368G19-HL		STS			481: ACTTACCAAGGATGGG	482: CAACCAAGGATGAG	
B368G19-HL		STS			483: TAGGCTCTGACTCTGG	484: ACCACAGGATCTCTG	
B368G19-HL		STS			485: TAAAGCGGTGAAGTGA	486: CTACCGCTCTCTAGGCT	
B368G19-HL		STS			487: TGGGCGCAGATATCTT	488: CTGGTGTGTGGTGTGT	
B368G19-HL		STS			489: AAGGAAGGATCAGAG	490: CACAAATTCATTTCCA	
B368G19-HL		STS			491: TCAATAGGATTCACATTT	492: AAGTCCCAAGAGGCTC	
B368G19-HL		STS			493: GGGTGAAGGATCTTTT	494: TGTGACATTCATGGC	
B368G19-HL		STS			495: GTCTGGGAAGATGGA	496: TCAAGGCTCTCTCTTAA	
B368G19-HL		STS			497: TCTTCTGCTACTTGG	498: TGGGAGGTGAGAGTGA	
B368G19-HL		STS			499: GGAAGTATGTTGGG	500: AGGAGCTGTTTGTGA	
B368G19-HL		STS			501: CTCTTGAAGTCCGCTGG	502: CAACGAGATCTCTAGC	
B368G19-HL		STS			503: GTGGGAGAGATCAAA	504: GTTTCAGAGAGAGCA	
B368G19-HL		STS			505: AGCTGTGAGGATCACT	506: GAGGATCTCTAGGAT	
B368G19-HL		STS			507: TAGGGGATCTTTTCCA	508: GAGCAATTTGAAAAGCCA	
B368G19-HL		STS			509: ATGGTCAAGCTCTCTGT	510: ATAGAGCACTCTCTCC	
B368G19-HL		STS			511: AATATGCTGTAGGCA	512: GAATTCGAAACAGATTC	
B368G19-HL		STS			513: ATGAGTTGGAGCTGGA	514: ATGAAAGTCTTGGCTCC	
B368G19-HL		STS			515: GAGGAGAGATCAAGAGG	516: TCTCTGGGAGTCTGAGCC	
B368G19-HL		STS			517: CTGAGCTTTTGGCTGT	518: CTCTAGGTGAGCAGAG	
B368G19-HL		STS			519: TGTAGCTCTGGAGGCTGT	520: ACACCTGCTGAGGAAAT	
B368G19-HL		STS			521: GAGGAGAGCTGATATA	522: TTTGCTTCTTACATGC	
B368G19-HL		STS			523: AAAATTTGAGCACTCC	524: TTTATTTTAAAGTGCTTTGTT	
B368G19-HL		STS			525: GTGCAAGCCACAGTAT	526: AGGAATTCGAGAGAG	
B368G19-HL		STS			527: CACTGATTTGATGATTTG	528: TCTGGTCTGAGTCTGCTA	
B368G19-HL		STS			529: AGATTTTGGGAGTCAAG	530: CGCTCAAGCAATCTC	
B368G19-HL		STS			531: GAGGCCCAAGTATGCA	532: GAATCATCAATCCAGCA	
B368G19-HL		STS			533: AGCTTCAAGTATGATCC	534: GAGGAGATCTGAGAG	
B368G19-HL		STS			535: ATGCTTTCAGCATTTGG	536: TGAATCTGTTAGGTTA	
B368G19-HL		STS			537: GTGGAATTTGTTCAAA	538: TTTATGGGAATTTGAGCC	
B368G19-HL		STS			539: TTTGGAAGAGAGAAATGT	540: GCTAGTCTTCTGAGCC	
B368G19-HL		STS			541: CTTTATGCTCTGATTC	542: GCTTTTCAAGCTGAGCA	
B368G19-HL		STS			543: TCAAGCTTCTTCTTCA	544: GTAGCCCAAGAGTGTCT	
B368G19-HL		STS			545: CTTGGCTGAGATAGGAT	546: CTCCCTCTGCTATGT	
B368G19-HL		STS			547: GGCAGTACTTCTTCCA	548: GTGCTCTCTTACAGCAA	
B368G19-HL		STS			549: ACCAGGCTGATGTGT	550: ACTGAGTATTAATATCTCCCT	
B368G19-HL		STS			551: GATGATTTTCTTCAAC	552: TCTGCTTTTAGAGCTTTAGC	
B368G19-HL		STS			553: TCAAGCTTCAAGAGAGCA	554: GAGGATATCCAGGACC	
B368G19-HL		STS			555: TGTGCTTTTAAATCAAGA	556: CTCCCTTACTTACTTGTG	
B368G19-HL		STS			557: TCTTCTCCAGGGAATCT	558: TTTATGCTCTGAGCAC	
B368G19-HL		STS			559: TCTTGGCTATCTTGAATC	560: TTTGCTGAGGCTGAGC	
B368G19-HL		STS			561: CAGAGAGCTGATATATCA	562: CATCTCTGCTCTTCTAGT	
B368G19-HL		STS			563: AGTTCAGAGAGAGAGAGC	564: CTGTCTCTCTCTGCTT	
B368G19-HL		STS			565: GAGGCTGAGAGGTTGAGAG	566: AAACAACTCCAGAGCACC	
B368G19-HL		STS			567: CTGAACACTTCTGATGCTG	568: CTAATCTTACTCTTACAGGCC	
B368G19-HL		STS			569: GAGGATTTTCAATCTTAACTG	570: CACTCTGATGAGCAATC	
B368G19-HL		STS			571: CTCTCTGCTGCTCTGAG	572: GGTGTCTTGAATCTGAGC	
B368G19-HL		STS			573: TGAGGATTAAGAGAGAG	574: TTTGTGCTCAATTTGAGTGG	
B368G19-HL		STS			575: AGGGAAGGATGCTTGG	576: TTTGCTGAGGAGGAGTGT	
B368G19-HL		STS			577: ATTGAGGCTCTTCAAGAGATGCTGAGC	578: AGAGGCTCAATCTTTTGGGAGAC	
B368G19-HL		STS			579: TTTGATTTGAGGATTTGCTGG	580: GGTGACATCTCTCTTCTC	
B368G19-HL		STS			581: TTTTGGCTCAATCTATGCC	582: GGTGAGAGAGAGAGCTCC	
B368G19-HL		STS			583: TTGCTCAAGTCTCTCTGG	584: ACCTTTTGTGAGGGAG	
B368G19-HL		STS			585: CTGGGATTTTGGAGAGC	586: GGGCAATTTACTACTTGC	

TABLE 7: HBM STS Table

STS Name	Locus Name	Type	GDB Access. #	Size (kb)	SEQ ID NO: Forward Primer	SEQ ID NO: Reverse Primer	Gene Name
B292118-HR		STS			587: CTTGTGTCAGTGTCTACGGG	588: TGGAAATTGTGTGTTCTTGG	
B10A18-HL		STS			589: CCAGTTCACCTGGATGTT	590: ATGGGCTGTGTGTTCTCAA	
B10A18-HR		STS			591: CTGGCTATCTCTGGACCTT	592: AGTTTGTCCCTAGTGGCCC	
B527D12-HL		STS			593: CAACAGTCTGACATCCAT	594: GGATAGTGGACACCA	
B372111-HR		STS			595: TGGGTGTGTAATTTGTTCAT	596: AGTTCCAGCCCTTACCAAG	
B372111-HL		STS			597: GGCCACTATCATCTCTGTGT	598: TTTCACATGGGAAGACAGG	
B37E17-HR(GS)		STS			599: ACAGTGACACTAGGAGAGGG	600: TCCAGGATGGAGATACAA	
B37E17-HL(GS)		STS			601: CTTGTGGCAGCATATCCAC	602: ACACCAAGAGATGGAGCCAC	
B34F22-HR(GS)		STS			603: TGTGTGTGTACCAAGTCCCA	604: TGAACGGAGGACCTACCAAG	
B34F22-HL(GS)		STS			605: GCAAGGTCCGACTCACTAAG	606: GCTGTGATTCCTTTACGC	
B64BP22-HR1		STS			607: ACAGTGGGGACAAAGACAGG	608: TACAGGGCACTCCAGTAG	
B82A4-HR2		STS			609: TCTTCTGTAAAGTTTCCCCC	610: TGTCTCAACCTCCCTCTGC	
B64BP22-HL		STS			611: AAGATATTTCTCCAGGCC	612: CAGTCCAGCCCATGAGAC	
B82L11-HL (GS)		STS			613: CTCTCTGATGGAGAGATC	614: AGACCTGGGACAGTCTGTG	
B86113-HL (GS)		STS			615: GGGAGACAGCTCACAGAT	616: TGAATTTGGAGATGGTGA	
144A24-HL		STS			617: CAGGCATCTCTATGTGCA	618: GGGAGGACAAAGTCTTTCA	
B82L11-HR (GS)		STS			619: ACTTGTGGCACTGAGTGTG	620: CCTTTCTTACGATGAGCA	
B86113-HR (GS)		STS			621: GGCCTGCTAGGCTCTTGAT	622: TGGGTCTCTCTGCTGACTT	
B82L11-HL2(GS)		STS			623: TCACCTACTTCCAGCTTCCG	624: AGACCTGGGACAGTCTGTG	
B82L11-HL3(GS)		STS			625: CTCCTCTGATGGAGATTC	626: AATTCAGGAGACCTGGGACC	

Novel STSs were developed either from publicly available genomic sequence or from sequence-derived BAC insert ends. Primers were chosen using a script which automatically performs vector and repetitive sequence masking using Cross_match (P. Green, Univ. of Washington) and subsequent primer picking using Primer3 (Rozen, Skaletsky (1996, 1997).
5 Primer3 is available at: www.genome.wi.mit.edu/genome_software/other/primer3.html.

Polymerase chain reaction (PCR) conditions for each primer pair were initially optimized with respect to MgCl₂ concentration. The standard buffer was 10 mM Tris-HCl (pH 8.3), 50 mM KCl, MgCl₂, 0.2 mM each dNTP, 0.2 µM each primer, 2.7 ng/µl human DNA, 0.25 units of AmpliTaq (Perkin Elmer) and MgCl₂ concentrations of 1.0 mM, 1.5 mM,
10 2.0 mM or 2.4 mM. Cycling conditions included an initial denaturation at 94°C for 2 minutes followed by 40 cycles at 94°C for 15 seconds, 55°C for 25 seconds, and 72°C for 25 seconds followed by a final extension at 72°C for 3 minutes. Depending on the results from the initial round of optimization the conditions were further optimized if necessary. Variables included increasing the annealing temperature to 58°C or 60°C, increasing the
15 cycle number to 42 and the annealing and extension times to 30 seconds, and using AmpliTaq™ Gold (Perkin Elmer).

BAC clones (Kim *et al.*, 1996 *Genomics* 32: 213-8; and Shizuya *et al.*, 1992 *Proc. Natl. Acad. Sci. USA* 89: 8794-7) containing STS markers of interest were obtained by PCR-based screening of DNA pools from a total human BAC library purchased from Research
20 Genetics. DNA pools derived from library plates 1-596 were used corresponding to nine genomic equivalents of human DNA. The initial screening process involved PCR reactions of individual markers against superpools, i.e., a mixture of DNA derived from all BAC clones from eight 384-well library plates. For each positive superpool, plate (8), row (16) and column (24), pools were screened to identify a unique library address. PCR products were
25 electrophoresed in 2% agarose gels (Sigma) containing 0.5 µg/ml ethidium bromide in 1X TBE at 150 volts for 45 min. The electrophoresis units used were the Model A3-1 systems from Owl Scientific Products. Typically, gels contained 10 tiers of lanes with 50 wells/tier. Molecular weight markers (100 bp ladder, Life Technologies, Bethesda, MD) were loaded at both ends of the gel. Images of the gels were captured with a Kodak DC40 CCD camera and
30 processed with Kodak 1D software. The gel data were exported as tab delimited text files;

names of the files included information about the library screened, the gel image files and the marker screened. These data were automatically imported using a customized Perl script into Filemaker™ PRO (Claris Corp.) databases for data storage and analysis. In cases where incomplete or ambiguous clone address information was obtained, additional experiments were performed to recover a unique, complete library address.

Recovery of clonal BAC cultures from the library involved streaking out a sample from the library well onto LB agar (Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982)) containing 12.5 µg/ml chloramphenicol (Sigma). Two individual colonies and a portion of the initial streak quadrant were tested with appropriate STS markers by colony PCR for verification. Positive clones were stored in LB broth containing 12.5 µg/ml chloramphenicol and 15% glycerol at -70°C.

Several different types of DNA preparation methods were used for isolation of BAC DNA. The manual alkaline lysis miniprep protocol listed below (Maniatis *et al.*, 1982) was successfully used for most applications, i.e., restriction mapping, CHEF gel analysis, FISH mapping, but was not successfully reproducible by endsequencing. The Autogen and Qiagen protocols were used specifically for BAC DNA preparation for endsequencing purposes.

Bacteria were grown in 15 ml Terrific Broth containing 12.5 µg/ml chloramphenicol in a 50 ml conical tube at 37°C for 20 hrs with shaking at 300 rpm. The cultures were centrifuged in a Sorvall RT 6000 D at 3000 rpm (~1800xg) at 4°C for 15 min. The supernatant was then aspirated as completely as possible. In some cases, cell pellets were frozen at -20°C at this step for up to 2 weeks. The pellet was then vortexed to homogenize the cells and minimize clumping. 250 µl of P1 solution (50 mM glucose, 15 mM Tris-HCl, pH 8, 10 mM EDTA, and 100 µg/ml RNase A) was added and the mixture pipetted up and down to mix. The mixture was then transferred to a 2 ml Eppendorf tube. 350 µl of P2 solution (0.2 N NaOH, 1% SDS) was then added, the mixture mixed gently and incubated for 5 min at room temperature. 350 µl of P3 solution (3 M KOAc, pH 5.5) was added and the mixture mixed gently until a white precipitate formed. The solution was incubated on ice for 5 min and then centrifuged at 4°C in a microfuge for 10 min. The supernatant was transferred carefully (avoiding the white precipitate) to a fresh 2 ml Eppendorf tube, and 0.9

ml of isopropanol was added, the solution mixed and left on ice for 5 min. The samples were centrifuged for 10 min, and the supernatant removed carefully. Pellets were washed in 70% ethanol and air dried for 5 min. Pellets were resuspended in 200 µl of TE8 (10 mM Tris-HCl, pH 8.0, 1.0 mM EDTA), and RNase A (Boehringer Mannheim) added to 100 µg/ml.

5 Samples were incubated at 37°C for 30 min, then precipitated by addition of $C_2H_3O_2Na \cdot 3H_2O$ to 0.5 M and 2 volumes of ethanol. Samples were centrifuged for 10 min, and the pellets washed with 70% ethanol followed by air drying and dissolving in 50 µl TE8. Typical yields for this DNA prep were 3-5 µg/15 ml bacterial culture. Ten to 15 µl were used for *HindIII* restriction analysis; 5 µl was used for *NotI* digestion and clone insert sizing by CHEF gel electrophoresis.

10 BACs were inoculated into 15 ml of 2X LB Broth containing 12.5 µg/ml chloramphenicol in a 50 ml conical tube. Four tubes were inoculated for each clone. Cultures were grown overnight (~16 hr) at 37°C with vigorous shaking (>300 rpm). Standard conditions for BAC DNA isolation were followed as recommended by the Autogen 15 740 manufacturer. 3 ml samples of culture were placed into Autogen tubes for a total of 60 ml or 20 tubes per clone. Samples were dissolved finally in 100 µl TE8 with 15 seconds of shaking as part of the Autogen protocol. After the Autogen protocol was finished DNA solutions were transferred from each individual tube and pooled into a 2 ml Eppendorf tube. Tubes with large amounts of debris (carry over from the pelleting debris step) were avoided. 20 The tubes were then rinsed with 0.5 ml of TE8 successively and this solution added to the pooled material. DNA solutions were stored at 4°C; clumping tended to occur upon freezing at -20°C. This DNA was either used directly for restriction mapping, CHEF gel analysis or FISH mapping or was further purified as described below for use in endsequencing reactions.

25 The volume of DNA solutions was adjusted to 2 ml with TE8, samples were then mixed gently and heated at 65°C for 10 min. The DNA solutions were then centrifuged at 4°C for 5 min and the supernatants transferred to a 15 ml conical tube. The NaCl concentration was then adjusted to 0.75 M (~0.3 ml of 5 M NaCl to the 2 ml sample). The total volume was then adjusted to 6 ml with Qiagen column equilibration buffer (Buffer QBT). The supernatant containing the DNA was then applied to the column and allowed to 30 enter by gravity flow. Columns were washed twice with 10 ml of Qiagen Buffer QC. Bound

DNA was then eluted with four separate 1 ml aliquots of Buffer QF kept at 65°C. DNA was precipitated with 0.7 volumes of isopropanol (~2.8 ml). Each sample was then transferred to 4 individual 2.2 ml Eppendorf tubes and incubated at room temperature for 2 hr or overnight. Samples were centrifuged in a microfuge for 10 min. at 4°C. The supernatant was removed carefully and 1 ml of 70% ethanol was added. Samples were centrifuged again and because the DNA pellets were often loose at this stage, the supernatant removed carefully. Samples were centrifuged again to concentrate remaining liquid which was removed with a micropipet tip. DNA pellets were then dried in a desiccator for 10 min. 20 µl of sterile distilled and deionized H₂O was added to each tube which was then placed at 4°C overnight. The four 20 µl samples for each clone were pooled and the tubes rinsed with another 20 µl of sterile distilled and deionized H₂O for a final volume of 100 µl. Samples were then heated at 65°C for 5 min. and then mixed gently. Typical yields were 2-5 µg/60 ml culture as assessed by *NotI* digestion and comparison with uncut lambda DNA.

3 ml of LB Broth containing 12.5 µg/ml of chloramphenicol was dispensed into autoclaved Autogen tubes. A single tube was used for each clone. For inoculation, glycerol stocks were removed from -70°C storage and placed on dry ice. A small portion of the glycerol stock was removed from the original tube with a sterile toothpick and transferred into the Autogen tube; the toothpick was left in the Autogen tube for at least two minutes before discarding. After inoculation the tubes were covered with tape making sure the seal was tight. When all samples were inoculated, the tube units were transferred into an Autogen rack holder and placed into a rotary shaker at 37°C for 16-17 hours at 250 rpm. Following growth, standard conditions for BAC DNA preparation, as defined by the manufacturer, were used to program the Autogen. Samples were not dissolved in TE8 as part of the program and DNA pellets were left dry. When the program was complete, the tubes were removed from the output tray and 30 µl of sterile distilled and deionized H₂O was added directly to the bottom of the tube. The tubes were then gently shaken for 2-5 seconds and then covered with parafilm and incubated at room temperature for 1-3 hours. DNA samples were then transferred to an Eppendorf tube and used either directly for sequencing or stored at 4°C for later use.

7. BAC Clone Characterization for Physical Mapping

DNA samples prepared either by manual alkaline lysis or the Autogen protocol were digested with *Hind*III for analysis of restriction fragment sizes. This data were used to compare the extent of overlap among clones. Typically 1-2 µg were used for each reaction. Reaction mixtures included: 1X Buffer 2 (New England Biolabs), 0.1 mg/ml bovine serum albumin (New England Biolabs), 50 µg/ml RNase A (Boehringer Mannheim), and 20 units of *Hind*III (New England Biolabs) in a final volume of 25 µl. Digestions were incubated at 37°C for 4-6 hours. BAC DNA was also digested with *Not*I for estimation of insert size by CHEF gel analysis (see below). Reaction conditions were identical to those for *Hind*III except that 20 units of *Not*I were used. Six µl of 6X Ficoll loading buffer containing bromphenol blue and xylene cyanol was added prior to electrophoresis.

*Hind*III digests were analyzed on 0.6% agarose (Seakem, FMC Bioproducts) in 1X TBE containing 0.5 µg/ml ethidium bromide. Gels (20 cm X 25 cm) were electrophoresed in a Model A4 electrophoresis unit (Owl Scientific) at 50 volts for 20-24 hrs. Molecular weight size markers included undigested lambda DNA, *Hind*III digested lambda DNA, and *Hae*III digested _X174 DNA. Molecular weight markers were heated at 65°C for 2 min prior to loading the gel. Images were captured with a Kodak DC40 CCD camera and analyzed with Kodak 1D software.

*Not*I digests were analyzed on a CHEF DRII (BioRad) electrophoresis unit according to the manufacturer's recommendations. Briefly, 1% agarose gels (BioRad pulsed field grade) were prepared in 0.5X TBE, equilibrated for 30 minutes in the electrophoresis unit at 14°C, and electrophoresed at 6 volts/cm for 14 hrs with circulation. Switching times were ramped from 10 sec to 20 sec. Gels were stained after electrophoresis in 0.5 µg/ml ethidium bromide. Molecular weight markers included undigested lambda DNA, *Hind*III digested lambda DNA, lambda ladder PFG ladder, and low range PFG marker (all from New England Biolabs).

BAC DNA prepared either by the manual alkaline lysis or Autogen protocols were labeled for FISH analysis using a Bioprime labeling kit (BioRad) according to the manufacturer's recommendation with minor modifications. Approximately 200 ng of DNA was used for each 50 µl reaction. 3 µl were analyzed on a 2% agarose gel to determine the

extent of labeling. Reactions were purified using a Sephadex G50 spin column prior to *in situ* hybridization. Metaphase FISH was performed as described (Ma *et al.*, 1996 *Cytogenet. Cell Genet.*, 74: 266-71).

8. **BAC Endsequencing**

The sequencing of BAC insert ends utilized DNA prepared by either of the two methods described above. The DYEnamic energy transfer primers and Dynamic Direct cycle sequencing kits from Amersham were used for sequencing reactions. Ready made sequencing mix including the M13 -40 forward sequencing primer was used (Catalog # US79730) for the T7 BAC vector terminus; ready made sequencing mix (Catalog # US79530) was mixed with the M13 -28 reverse sequencing primer (Catalog # US79339) for the SP6 BAC vector terminus. The sequencing reaction mixes included one of the four fluorescently labeled dye-primers, one of the four dideoxy termination mixes, dNTPs, reaction buffer, and Thermosequenase. For each BAC DNA sample, 3 μ l of the BAC DNA sample was aliquoted to 4 PCR strip tubes. 2 μ l of one of the four dye primer/termination mix combinations was then added to each of the four tubes. The tubes were then sealed and centrifuged briefly prior to PCR. Thermocycling conditions involved a 1 minute denaturation at 95°C, 15 second annealing at 45°C, and extension for 1 minute at 70°C for 35 total cycles. After cycling the plates were centrifuged briefly to collect all the liquid to the bottom of the tubes. 5 μ l of sterile distilled and deionized H₂O was then added into each tube, the plates sealed and centrifuged briefly again. The four samples for each BAC were then pooled together. DNA was then precipitated by adding 1.5 μ l of 7.5 M NH₄OAc and 100 μ l of -20°C 100% ethanol to each tube. Samples were mixed by pipetting up and down once. The plates were then sealed and incubated on ice for 10 minutes. Plates were centrifuged in a table top Haraeus centrifuge at 4000 rpm (3,290xg) for 30 minutes at 4°C to recover the DNA. The supernatant was removed and excess liquid blotted onto paper towels. Pellets were washed by adding 100 μ l of -20°C 70% ethanol into each tube and re-centrifuging at 4000 rpm (3,290xg) for 10 minutes at 4°C. The supernatant was removed and excess liquid again removed by blotting on a paper towel. Remaining traces of liquid were removed by placing the plates upside down over a paper towel and centrifuging only until the centrifuge

reached 800 rpm. Samples were then air dried at room temperature for 30 min. Tubes were capped and stored dry at -20°C until electrophoresis. Immediately prior to electrophoresis the DNA was dissolved in 1.5 µl of Amersham loading dye. Plates were then sealed and centrifuged at 2000 rpm (825Xg). The plates were then vortexed on a plate shaker for 1-2 minutes. Samples were then recentrifuged at 2000 rpm (825Xg) briefly. Samples were then heated at 65°C for 2 min and immediately placed on ice. Standard gel electrophoresis was performed on ABI 377 fluorescent sequencers according to the manufacturer's recommendation.

9. Sub-cloning and Sequencing of HBM BAC DNA

The physical map of the *Zmax1* (*LRP5*) gene region provides a set of BAC clones that contain within them the *Zmax1* (*LRP5*) gene and the *HBM* gene. DNA sequencing of several of the BACs from the region has been completed. The DNA sequence data is a unique reagent that includes data that one skilled in the art can use to identify the *Zmax1* (*LRP5*) gene and the *HBM* gene, or to prepare probes to identify the gene(s), or to identify DNA sequence polymorphisms that identify the gene(s).

BAC DNA was isolated according to one of two protocols, either a Qiagen purification of BAC DNA (Qiagen, Inc. as described in the product literature) or a manual purification which is a modification of the standard alkaline lysis/Cesium Chloride preparation of plasmid DNA (see e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). Briefly for the manual protocol, cells were pelleted, resuspended in GTE (50 mM glucose, 25 mM Tris-Cl (pH 8), 10 mM EDTA) and lysozyme (50 mg/ml solution), followed by NaOH/SDS (1% SDS/0.2 N NaOH) and then an ice-cold solution of 3 M KOAc (pH 4.5-4.8). RNaseA was added to the filtered supernatant, followed by Proteinase K and 20% SDS. The DNA was then precipitated with isopropanol, dried and resuspended in TE (10 mM Tris, 1 mM EDTA (pH 8.0)). The BAC DNA was further purified by Cesium Chloride density gradient centrifugation (Ausubel *et al.*, 1997).

Following isolation, the BAC DNA was sheared hydrodynamically using an HPLC (Hengen, 1997 *Trends in Biochem. Sci.* 22: 273-4) to an insert size of 2000-3000 bp. After shearing, the DNA was concentrated and separated on a standard 1% agarose gel. A single

fraction, corresponding to the approximate size, was excised from the gel and purified by electroelution (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring, NY (1989)).

The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The blunt-ended DNA was then ligated to unique *Bst*XI-linker adapters (5'- GTCTTCACCACG GGG-3' and 5'-GTGGTGAAGAC-3' in 100-1000 fold molar excess; SEQ ID NOS: 627 and 628 respectively). These linkers were complimentary to the *Bst*XI-cut pMPX vectors (constructed by the inventors), while the overhang was not self-complimentary. Therefore, the linkers would not concatemerize nor would the cut-vector religate itself easily. The linker-adapted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean (BIO 101, Inc.). The linker-adapted insert was then ligated to a modified pBlueScript vector to construct a "shotgun" subclone library. The vector contained an out-of-frame lacZ gene at the cloning site which became in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

All subsequent steps were based on sequencing by ABI377 automated DNA sequencing methods. Only major modifications to the protocols are highlighted. Briefly, the library was then transformed into DH5 α competent cells (Life Technologies, Bethesda, MD, DH5 α transformation protocol). It was assessed by plating onto antibiotic plates containing ampicillin and IPTG/Xgal. The plates were incubated overnight at 37°C. Successful transformants were then used for plating of clones and picking for sequencing. The cultures were grown overnight at 37°. DNA was purified using a silica bead DNA preparation (Ng *et al.*, 1996 *Nucl. Acids Res.* 24: 5045-7) method. In this manner, 25 μ g of DNA was obtained per clone.

These purified DNA samples were then sequenced using ABI dye-terminator chemistry. The ABI dye terminator sequence reads were run on ABI377 machines and the data was directly transferred to UNIX machines following lane tracking of the gels. All reads were assembled using PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, Jan. 1996, p.157) with default parameters and quality scores. The initial assembly was done at 6-fold coverage and yielded an average of 8-15 contigs. Following the initial assembly, missing mates (sequences from clones that only

gave one strand reads) were identified and sequenced with ABI technology to allow the identification of additional overlapping contigs. Primers for walking were selected using a Genome Therapeutics program Pick_primer near the ends of the clones to facilitate gap closure. These walks were sequenced using the selected clones and primers. Data were
5 reassembled with PHRAP into sequence contigs.

10. Gene Identification by Computational Methods

Following assembly of the BAC sequences into contigs, the contigs were subjected to computational analyses to identify coding regions and regions bearing DNA sequence
10 similarity to known genes. This protocol included the following steps.

1. Degap the contigs: the sequence contigs often contain symbols (denoted by a period symbol) that represent locations where the individual ABI sequence reads have insertions or deletions. Prior to automated computational analysis of the contigs, the periods were removed. The original data was maintained
15 for future reference.

2. BAC vector sequences were "masked" within the sequence by using the program cross match (Phil Green, <http://chimera.biotech.washington.edu/UWGC>). Since the shotgun libraries construction detailed above leaves some BAC vector in the shotgun libraries, this program was used to compare the sequence of the BAC contigs
20 to the BAC vector and to mask any vector sequence prior to subsequent steps. Masked sequences were marked by an "X" in the sequence files, and remained inert during subsequent analyses.

3. *E. coli* sequences contaminating the BAC sequences were masked by comparing the BAC contigs to the entire *E. coli* DNA sequence.

25 4. Repetitive elements known to be common in the human genome were masked using cross match. In this implementation of crossmatch, the BAC sequence was compared to a database of human repetitive elements (Jerzy Jerka, Genetic Information Research Institute, Palo Alto, CA). The masked repeats were marked by X and remained inert during subsequent analyses.

30 5. The location of exons within the sequence was predicted using the

MZEF computer program (Zhang, 1997 *Proc. Natl. Acad. Sci USA*. 94: 565-8).

6. The sequence was compared to the publicly available unigene database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using the blastn2 algorithm (Altschul *et al.*, 1997 *Nucl. Acids Res.* 25: 3389-3402). The parameters for this search were: E=0.05, v=50, B=50 (where E is the expected probability score cutoff, V is the number of database entries returned in the reporting of the results, and B is the number of sequence alignments returned in the reporting of the results (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)).

7. The sequence was translated into protein for all six reading frames, and the protein sequences were compared to a non-redundant protein database compiled from Genpept Swissprot PIR (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov). The parameters for this search were E=0.05, V=50, B= 50, where E, V, and B are defined as above.

8. The BAC DNA sequence was compared to the database of the cDNA clones derived from direct selection experiments (described below) using blastn2 (Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=250, B=250, where E, V, and B are defined as above.

9. The BAC sequence was compared to the sequences of all other BACs from the HBM region on chromosome 11q12-13 using blastn2 (Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=50, B=50, where E, V, and B are defined as above.

10. The BAC sequence was compared to the sequences derived from the ends of BACs from the HBM region on chromosome 11q12-13 using blastn2 (Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=50, B=50, where E, V, and B are defined as above.

11. The BAC sequence was compared to the GenBank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using blastn2

(Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=50, B=50, where E, V, and B are defined as above.

12. The BAC sequence was compared to the STS division of GenBank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using blastn2 (Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=50, B= 50, where E, V, and B are defined as above.

13. The BAC sequence was compared to the Expressed Sequence (EST) Tag GenBank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using blastn2 (Altschul *et al.*, 1997). The parameters for this search were E=0.05, V=250, B=250, where E, V, and B are defined as above.

11. Gene Identification by Direct cDNA Selection

Primary linkered cDNA pools were prepared from bone marrow, calvarial bone, femoral bone, kidney, skeletal muscle, testis and total brain. Poly (A) + RNA was prepared from calvarial and femoral bone tissue (Chomczynski *et al.*, 1987 *Anal. Biochem.* 162: 156-9; and D'Alessio *et al.*, 1987 *Focus* 9: 1-4) and the remainder of the mRNA was purchased from Clontech (Palo Alto, California). In order to generate oligo(dT) and random primed cDNA pools from the same tissue, 2.5 µg mRNA was mixed with oligo(dT) primer in one reaction and 2.5 µg mRNA was mixed with random hexamers in another reaction, and both were converted to first and second strand cDNA according to manufacturers recommendations (Life Technologies, Bethesda, MD). Paired phosphorylated cDNA linkers (see sequence below) were annealed together by mixing in a 1:1 ratio (10 µg each) incubated at 65°C for five minutes and allowed to cool to room temperature.

Paired linkers oligo ½

OLIGO 1: 5'-CTG AGC GGA ATT CGT GAG ACC-3' (SEQ ID NO:12)

OLIGO 2: 5'-TTG GTC TCA CGT ATT CCG CTC GA-3' (SEQ ID NO:13)

Paired linkers oligo3/4

OLIGO 3: 5'-CTC GAG AAT TCT GGA TCC TC-3' (SEQ ID NO:14)

OLIGO 4: 5'-TTG AGG ATC CAG AAT TCT CGA G-3' (SEQ ID NO:15)

Paired linkers oligo5/6

OLIGO 5: 5'-TGT ATG CGA ATT CGC TGC GCG-3' (SEQ ID NO:16)

OLIGO 6: 5'-TTC GCG CAG CGA ATT CGC ATA CA-3' (SEQ ID NO:17)

5 Paired linkers oligo7/8

OLIGO 7: 5'-GTC CAC TGA ATT CTC AGT GAG-3' (SEQ ID NO:18)

OLIGO 8: 5'-TTG TCA CTG AGA ATT CAG TGG AC-3' (SEQ ID NO:19)

Paired linkers oligo11/12

OLIGO 11: 5'-GAA TCC GAA TTC CTG GTC AGC-3' (SEQ ID NO:20)

10 OLIGO 12: 5'-TTG CTG ACC AGG AAT TCG GAT TC-3' (SEQ ID NO:21)

Linkers were ligated to all oligo(dT) and random primed cDNA pools (see below) according to manufacturers instructions (Life Technologies, Bethesda, MD).

Oligo 1/2 was ligated to oligo(dT) and random primed cDNA pools prepared from bone marrow. Oligo 3/4 was ligated to oligo(dT) and random primed cDNA pools prepared from calvarial bone. Oligo 5/6 was ligated to oligo(dT) and random primed cDNA pools prepared from brain and skeletal muscle. Oligo 7/8 was ligated to oligo(dT) and random primed cDNA pools prepared from kidney. Oligo 11/12 was ligated to oligo(dT) and random primed cDNA pools prepared from femoral bone.

The cDNA pools were evaluated for length distribution by PCR amplification using 1 µl of a 1:1, 1:10, and 1:100 dilution of the ligation reaction, respectively. PCR reactions were performed in a Perkin Elmer 9600, each 25 µl volume reaction contained 1 µl of DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.001% gelatin, 200 mM each dNTPs, 10 µM primer and 1 unit Taq DNA polymerase (Perkin Elmer) and was amplified under the following conditions: 30 seconds at 94°C, 30 seconds at 60°C and 2 minutes at 72°C for 30 cycles. The length distribution of the amplified cDNA pools were evaluated by electrophoresis on a 1% agarose gel. The PCR reaction that gave the best representation of the random primed and oligo(dT) primed cDNA pools was scaled up so that ~2-3 µg of each cDNA pool was produced. The starting cDNA for the direct selection reaction comprised of 0.5 µg of random primed cDNAs mixed with 0.5 µg of oligo(dT) primed cDNAs.

30 The DNA from the 54 BACs that were used in the direct cDNA selection procedure

was isolated using Nucleobond AX columns as described by the manufacturer (The Nest Group, Inc.).

The BACs were pooled in equimolar amounts and 1 µg of the isolated genomic DNA was labeled with biotin 16-UTP by nick translation in accordance with the manufacturers instructions (Boehringer Mannheim). The incorporation of the biotin was monitored by methods that could be practiced by one skilled in the art (Del Mastro *et al.*, *Methods in Molecular Biology*, Humana Press Inc., NJ (1996)).

Direct cDNA selection was performed using methods that could be practiced by one skilled in the art (Del Mastro *et al.*, 1996). Briefly, the cDNA pools were multiplexed in two separate reactions: In one reaction cDNA pools from bone marrow, calvarial bone, brain and testis were mixed, and in the other cDNA pools from skeletal muscle, kidney and femoral bone were mixed. Suppression of the repeats, yeast sequences and plasmid in the cDNA pools was performed to a Cot of 20. 100 ng of biotinylated BAC DNA was mixed with the suppressed cDNAs and hybridized in solution to a Cot of 200. The biotinylated DNA and the cognate cDNAs was captured on streptavidin-coated paramagnetic beads. The beads were washed and the primary selected cDNAs were eluted. These cDNAs were PCR amplified and a second round of direct selection was performed. The product of the second round of direct selection is referred to as the secondary selected material. A Galanin cDNA clone, previously shown to map to 11q12-13 (Evans, 1993 *Genomics* 18: 473-7), was used to monitor enrichment during the two rounds of selection.

The secondary selected material from bone marrow, calvarial bone, femoral bone, kidney, skeletal muscle, testis and total brain was PCR amplified using modified primers of oligos 1, 3, 5, 7 and 11, shown below, and cloned into the UDG vector pAMP10 (Life Technologies, Bethesda, MD), in accordance with the manufacturer's recommendations.

Modified primer sequences:

Primer	SEQ ID NO.	Sequence
Oligo 1	22	5'-CUA CUA CUA CUA CTG AGC GGA ATT CGT GAG ACC-3'
Oligo 3	23	5'-CUA CUA CUA CUA CTC GAG AAT TCT GGA TCC TC-3'
Oligo 5	24	5'-CUA CUA CUA CUA TGT ATG CGA ATT CGC TGC GCG-3'

Primer	SEQ ID NO.	Sequence
Oligo 7	25	5'-CUA CUA CUA CUA GTC CAC TGA ATT CTC AGT GAG-3'
Oligo 11	26	5'-CUA CUA CUA CUA GAA TCC GAA TTC CTG GTC AGC-3'

The cloned secondary selected material, from each tissue source, was transformed into
 5 MAX Efficiency DH5a Competent Cells (Life Technologies, Bethesda, MD) as
 recommended by the manufacturer. 384 colonies were picked from each transformed source
 and arrayed into four 96 well microtiter plates.

All secondarily selected cDNA clones were sequenced using M13 dye primer
 terminator cycle sequencing kit (Applied Biosystems), and the data collected by the ABI 377
 10 automated fluorescence sequencer (Applied Biosystems).

All sequences were analyzed using the BLASTN, BLASTX and FASTA programs
 (Altschul *et al.*, 1990 *J. Mol. Biol.* 215: 403-410; and Altschul *et al.*, *Nucl. Acids. Res.* 25:
 3389-3402). The cDNA sequences were compared to a database containing sequences
 derived from human repeats, mitochondrial DNA, ribosomal RNA, *E. coli* DNA to remove
 15 background clones from the dataset using the program cross_match. A further round of
 comparison was also performed using the program BLASTN2 against known genes
 (GenBank) and the BAC sequences from the HBM region. Those cDNAs that were >90%
 homologous to these sequences were filed according to the result and the data stored in a
 database for further analysis. cDNA sequences that were identified but did not have
 20 significant similarity to the BAC sequences from the HBM region or were eliminated by
 cross_match were hybridized to nylon membranes which contained the BACs from the HBM
 region, to ascertain whether they hybridized to the target.

Hybridization analysis was used to map the cDNA clones to the BAC target that
 selected them. The BACs that were identified from the HBM region were arrayed and grown
 25 into a 96 well microtiter plate. LB agar containing 25 µg/ml kanamycin was poured into 96
 well microtiter plate lids. Once the agar had solidified, pre-cut Hybond N+ nylon membranes
 (Amersham) were laid on top of the agar and the BACs were stamped onto the membranes in
 duplicate using a hand held 96 well replica plater (V&P Scientific, Inc.). The plates were

incubated overnight at 37°C. The membranes were processed according to the manufacturers recommendations.

The cDNAs that needed to be mapped by hybridization were PCR amplified using the relevant primer (oligos 1, 3, 5, 7 and 11) that would amplify that clone. For this PCR
5 amplification, the primers were modified to contain a linkered digoxigenin molecule at the 5' of the oligonucleotide. The PCR amplification was performed under the same conditions as described in Preparation of cDNA Pools (above). The PCR products were evaluated for quality and quantity by electrophoresis on a 1% agarose gel by loading 5 µl of the PCR reaction. The nylon membranes containing the stamped BACs were individually pre-
10 hybridized in 50 ml conical tubes containing 10 ml of hybridization solution (5X SSPE, 0.5X Blotto, 2.5% SDS and 1 mM EDTA (pH 8.0)). The 50 ml conical tubes were placed in a rotisserie oven (Robbins Scientific) for 2 hours at 65°C. Twenty-five ng of each cDNA probe was denatured and added into individual 50 ml conical tubes containing the nylon membrane and hybridization solution. The hybridization was performed overnight at 65°C.
15 The filters were washed for 20 minutes at 65°C in each of the following solutions: 3X SSPE, 0.1% SDS; 1X SSPE, 0.1% SDS and 0.1X SSPE, 0.1% SDS.

The membranes were removed from the 50 ml conical tubes and placed in a dish. Acetate sheets were placed between each membrane to prevent them from sticking to each other. The incubation of the membranes with the Anti-DIG-AP and CDP-Star was performed
20 according to manufacturers recommendations (Boehringer Mannheim). The membranes were wrapped in Saran wrap and exposed to Kodak Bio-Max X-ray film for 1 hour.

12. cDNA Cloning and Expression Analysis

To characterize the expression of the genes identified by direct cDNA selection and
25 genomic DNA sequencing in comparison to the publicly available databases, a series of experiments were performed to further characterize the genes in the HBM region. First, oligonucleotide primers were designed for use in the polymerase chain reaction (PCR) so that portions of a cDNA, EST, or genomic DNA could be amplified from a pool of DNA molecules (a cDNA library) or RNA population (RT-PCR and RACE). The PCR primers
30 were used in a reaction containing genomic DNA to verify that they generated a product of

the size predicted based on the genomic (BAC) sequence. A number of cDNA libraries were then examined for the presence of the specific cDNA or EST. The presence of a fragment of a transcription unit in a particular cDNA library indicates a high probability that additional portions of the same transcription unit will be present as well.

5 A critical piece of data that is required when characterizing novel genes is the length, in nucleotides, of the processed transcript or messenger RNA (mRNA). One skilled in the art primarily determines the length of an mRNA by Northern blot hybridization (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). Groups of ESTs and direct-selected cDNA clones that displayed
10 significant sequence similarity to sequenced BACs in the critical region were grouped for convenience into approximately 30 kilobase units. Within each 30 kilobase unit there were from one up to fifty ESTs and direct-selected cDNA clones which comprised one or more independent transcription units. One or more ESTs or direct-selected cDNAs were used as hybridization probes to determine the length of the mRNA in a variety of tissues, using
15 commercially available reagents (Multiple Tissue Northern blot; Clontech, Palo Alto, California) under conditions recommended by the manufacturer.

Directionally cloned cDNA libraries from femoral bone, and calvarial bone tissue were constructed by methods familiar to one skilled in the art (for example, Soares in *Automated DNA Sequencing and Analysis*, Adams, Fields and Venter, Eds., Academic Press,
20 NY, pages 110-114 (1994)). Bones were initially broken into fragments with a hammer, and the small pieces were frozen in liquid nitrogen and reduced to a powder in a tissue pulverizer (Spectrum Laboratory Products). RNA was extracted from the powdered bone by homogenizing the powdered bone with a standard Acid Guanidinium Thiocyanate-Phenol-Chloroform extraction buffer (e.g., Chomczynski *et al.*, 1987 *Anal.*
25 *Biochem.* 162: 156-9) using a polytron homogenizer (Brinkman Instruments). Additionally, human brain and lung total RNA was purchased from Clontech. PolyA RNA was isolated from total RNA using dynabeads-dT according to the manufacturer's recommendations (Dynal, Inc.). First strand cDNA synthesis was initiated using an oligonucleotide primer with the sequence: 5'-AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT
30 TT-3' (SEQ ID NO:27). This primer introduces a *NotI* restriction site (underlined) at the 3'

end of the cDNA. First and second strand synthesis were performed using the "one-tube" cDNA synthesis kit as described by the manufacturer (Life Technologies, Bethesda, MD). Double stranded cDNAs were treated with T4 polynucleotide kinase to ensure that the ends of the molecules were blunt (Soares in *Automated DNA Sequencing and Analysis*, Adams, Fields and Venter, Eds., Academic Press, NY, pages 110-114 (1994)), and the blunt ended cDNAs were then size selected by a Biogel column (Huynh *et al.* in *DNA Cloning*, Vol. 1, Glover, Ed., IRL Press, Oxford, pages 49-78 (1985)) or with a size-sep 400 Sepharose column (Pharmacia, catalog # 27-5105-01). Only cDNAs of 400 base pairs or longer were used in subsequent steps. *EcoRI* adapters (sequence: 5' OH-AATTCGGCACGAG-OH 3' (SEQ ID NO:28), and 5' p-CTCGTGCCG-OH 3' (SEQ ID NO:29)) were then ligated to the double stranded cDNAs by methods familiar to one skilled in the art (Soares, 1994). The *EcoRI* adapters were then removed from the 3' end of the cDNA by digestion with *NotI* (Soares, 1994). The cDNA was then ligated into the plasmid vector pBluescript® II KS+ (Stratagene, La Jolla, California), and the ligated material was transformed into *E. coli* host DH10B or DH12S by electroporation methods familiar to one skilled in the art (Soares, 1994). After growth overnight at 37°C, DNA was recovered from the *E. coli* colonies after scraping the plates by processing as directed for the Mega-prep kit (Qiagen, Chatsworth, California). The quality of the cDNA libraries was estimated by counting a portion of the total numbers of primary transformants and determining the average insert size and the percentage of plasmids with no cDNA insert. Additional cDNA libraries (human total brain, heart, kidney, leukocyte, and fetal brain) were purchased from Life Technologies, Bethesda, MD.

cDNA libraries, both oligo (dT) and random hexamer (N_6) primed, were used for isolating cDNA clones transcribed within the HBM region: human bone, human brain, human kidney and human skeletal muscle (all cDNA libraries were made by the inventors, except for skeletal muscle (dT) and kidney (dT) cDNA libraries). Four 10 x 10 arrays of each of the cDNA libraries were prepared as follows: the cDNA libraries were titrated to 2.5×10^6 using primary transformants. The appropriate volume of frozen stock was used to inoculate 2 L of LB/ampicillin (100 mg/ml). This inoculated liquid culture was aliquoted into 400 tubes of 4 ml each. Each tube contained approximately 5000 cfu. The tubes were incubated at 30°C

overnight with gentle agitation. The cultures were grown to an OD of 0.7-0.9. Frozen stocks were prepared for each of the cultures by allocating 100 µl of culture and 300 µl of 80% glycerol. Stocks were frozen in a dry ice/ethanol bath and stored at -70°C. The remaining culture was DNA prepared using the Qiagen (Chatsworth, CA) spin miniprep kit according to the manufacturer's instructions. The DNAs from the 400 cultures were pooled to make 80 column and row pools. The cDNA libraries were determined to contain HBM cDNA clones of interest by PCR. Markers were designed to amplify putative exons. Once a standard PCR optimization was performed and specific cDNA libraries were determined to contain cDNA clones of interest, the markers were used to screen the arrayed library. Positive addresses indicating the presence of cDNA clones were confirmed by a second PCR using the same markers.

Once a cDNA library was identified as likely to contain cDNA clones corresponding to a specific transcript of interest from the HBM region, it was manipulated to isolate the clone or clones containing cDNA inserts identical to the EST or direct-selected cDNA of interest. This was accomplished by a modification of the standard "colony screening" method (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). Specifically, twenty 150 mm LB+ampicillin agar plates were spread with 20,000 colony forming units (cfu) of cDNA library and the colonies allowed to grow overnight at 37°C. Colonies were transferred to nylon filters (Hybond from Amersham, or equivalent) and duplicates prepared by pressing two filters together essentially as described (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). The "master" plate was then incubated an additional 6-8 hours to allow the colonies to grow back. The DNA from the bacterial colonies was then affixed to the nylon filters by treating the filters sequentially with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for two minutes, neutralization solution (0.5 M Tris-Cl pH 8.0, 1.5 M NaCl) for two minutes (twice). The bacterial colonies were removed from the filters by washing in a solution of 2X SSC/0.1% SDS for one minute while rubbing with tissue paper. The filters were air dried and baked under vacuum at 80°C for 1-2 hours.

A cDNA hybridization probe was prepared by random hexamer labeling (Fineberg

and Vogelstein, *Anal. Biochem.*, 132:6-13 (1983)) or by including gene-specific primers and no random hexamers in the reaction (for small fragments). Specific activity was calculated and was $>5 \times 10^8$ cpm/ 10^8 μ g of cDNA. The colony membranes were then prewashed in 10 mM Tris-Cl pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS for 30 minutes at 55°C. Following
5 the prewash, the filters were prehybridized in > 2 ml/filter of 6X SSC, 50 % deionized formamide, 2% SDS, 5X Denhardt's solution, and 100 mg/ml denatured salmon sperm DNA, at 42°C for 30 minutes. The filters were then transferred to hybridization solution (6X SSC, 2% SDS, 5X Denhardt's, 100 mg/ml denatured salmon sperm DNA) containing denatured α^{32} P-dCTP-labeled cDNA probe and incubated at 42°C for 16-18 hours.

10 After the 16-18 hour incubation, the filters were washed under constant agitation in 2X SSC, 2% SDS at room temperature for 20 minutes, followed by two washes at 65°C for 15 minutes each. A second wash was performed in 0.5 X SSC, 0.5% SDS for 15 minutes at 65°C. Filters were then wrapped in plastic wrap and exposed to radiographic film for several hours to overnight. After film development, individual colonies on plates were aligned with
15 the autoradiograph so that they could be picked into a 1 ml solution of LB Broth containing ampicillin. After shaking at 37°C for 1-2 hours, aliquots of the solution were plated on 150 mm plates for secondary screening. Secondary screening was identical to primary screening (above) except that it was performed on plates containing ~250 colonies so that individual colonies could be clearly identified for picking.

20 After colony screening with radiolabeled probes yielded cDNA clones, the clones were characterized by restriction endonuclease cleavage, PCR, and direct sequencing to confirm the sequence identity between the original probe and the isolated clone. To obtain the full-length cDNA, the novel sequence from the end of the clone identified was used to probe the library again. This process was repeated until the length of the cDNA cloned
25 matches that estimated to be full-length by the northern blot analysis.

RT-PCR was used as another method to isolate full length clones. The cDNA was synthesized and amplified using a "Superscript One Step RT-PCR" kit (Life Technologies, Gaithersburg, MD). The procedure involved adding 1.5 μ g of RNA to the following: 25 μ l of reaction mix provided which is a proprietary buffer mix with MgSO_4 and dNTP's, 1 μ l sense
30 primer (10 μ M) and 1 μ l anti-sense primer (10 μ M), 1 μ l reverse transcriptase and *Taq* DNA

polymerase mix provided and autoclaved water to a total reaction mix of 50 μ l. The reaction was then placed in a thermocycler for 1 cycle at 50°C for 15 to 30 minutes, then 94°C for 15 seconds, 55-60°C for 30 seconds and 68-72°C for 1 minute per kilobase of anticipated product and finally 1 cycle of 72°C for 5-10 minutes. The sample was analyzed on an agarose gel. The product was excised from the gel and purified from the gel (GeneClean, Bio 101). The purified product was cloned in pCTNR (General Contractor DNA Cloning System, 5 Prime - 3 Prime, Inc.) and sequenced to verify that the clone was specific to the gene of interest.

Rapid Amplification of cDNA ends (RACE) was performed following the manufacturer's instructions using a Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA) as a method for cloning the 5' and 3' ends of candidate genes. cDNA pools were prepared from total RNA by performing first strand synthesis, where a sample of total RNA sample was mixed with a modified oligo (dT) primer, heated to 70°C, cooled on ice and followed by the addition of: 5X first strand buffer, 10 mM dNTP mix, and AMV Reverse Transcriptase (20 U/ μ l). The tube was incubated at 42°C for one hour and then the reaction tube was placed on ice. For second strand synthesis, the following components were added directly to the reaction tube: 5x second strand buffer, 10 mM dNTP mix, sterile water, 20X second strand enzyme cocktail and the reaction tube was incubated at 16°C for 1.5 hours. T4 DNA Polymerase was added to the reaction tube and incubated at 16°C for 45 minutes. The second-strand synthesis was terminated with the addition of an EDTA/Glycogen mix. The sample was subjected to a phenol/chloroform extraction and an ammonium acetate precipitation. The cDNA pools were checked for quality by analyzing on an agarose gel for size distribution. Marathon cDNA adapters (Clontech) were then ligated onto the cDNA ends. The specific adapters contained priming sites that allowed for amplification of either 5' or 3' ends, depending on the orientation of the gene specific primer (GSP) that was chosen. An aliquot of the double stranded cDNA was added to the following reagents: 10 μ M Marathon cDNA adapter, 5x DNA ligation buffer, T4 DNA ligase. The reaction was incubated at 16°C overnight. The reaction was heat inactivated to terminate the reaction. PCR was performed by the addition of the following to the diluted double stranded cDNA pool: 10X cDNA PCR reaction buffer, 10 μ M dNTP mix, 10 μ M GSP, 10 μ M AP1 primer

(kit), 50x Advantage cDNA Polymerase Mix. Thermal Cycling conditions were 94°C for 30 seconds, 5 cycles of 94°C for 5 seconds, 72°C for 4 minutes, 5 cycles of 94°C for 5 seconds, 70°C for 4 minutes, 23 cycles of 94°C for 5 seconds, 68°C for 4 minutes. After the first round of PCR was performed using the GSP to extend to the end of the adapter to create the adapter primer binding site, exponential amplification of the specific cDNA of interest was observed. Usually a second nested PCR is performed to confirm the specific cDNA. The RACE product was analyzed on an agarose gel and then excised and purified from the gel (GeneClean, BIO 101). The RACE product was then cloned into pCTNR (General Contractor DNA Cloning System, 5' - 3', Inc.) and the DNA sequence determined to verify that the clone is specific to the gene of interest.

13. Mutation Analysis

Comparative genes were identified using the above procedures and the exons from each gene were subjected to mutation detection analysis. Comparative DNA sequencing was used to identify polymorphisms in HBM candidate genes from chromosome 11q12-13. DNA sequences for candidate genes were amplified from patient lymphoblastoid cell lines.

The inventors developed a method based on analysis of direct DNA sequencing of PCR products amplified from candidate regions to search for the causative polymorphism. The procedure consisted of three stages that used different subsets of HBM family to find segregating polymorphisms and a population panel to assess the frequency of the polymorphisms. The family resources result from a single founder leading to the assumption that all affected individuals will share the same causative polymorphism.

Candidate regions were first screened in a subset of the HBM family consisting of the proband, daughter, and her mother, father and brother. Monochromosomal reference sequences were produced concurrently and used for comparison. The mother and daughter carried the HBM polymorphism in this nuclear family, providing the ability to monitor polymorphism transmission. The net result is that two HBM chromosomes and six non-HBM chromosomes were screened. This allowed exclusion of numerous frequent alleles. Only alleles exclusively present in the affected individuals passed to the next level of analysis.

Polymorphisms that segregated exclusively with the HBM phenotype in this original family were then re-examined in an extended portion of the HBM pedigree consisting of two additional nuclear families. These families consisted of five HBM and three unaffected individuals. The HBM individuals in this group included the two critical crossover
5 individuals, providing the centromeric and telomeric boundaries of the critical region. Tracking the heredity of polymorphisms between these individuals and their affected parents allowed for further refining of the critical region. This group brought the total of HBM chromosomes screened to seven and the total of non-HBM chromosomes to seventeen.

When a given polymorphism continued to segregate exclusively with the HBM
10 phenotype in the extended group, a population panel was then examined. This panel of 84 persons consisted of 42 individuals known to have normal bone mineral density and 42 individuals known to be unrelated but with untyped bone mineral density. For this purpose, normal bone mineral density is within two standard deviations of a BMD Z score of 0. The second group was from the widely used CEPH panel of individuals. Any segregating
15 polymorphisms found to be rare in this population were subsequently examined on the entire HBM pedigree and a larger population.

Polymerase chain reaction (PCR) was used to generate sequencing templates from the HBM family's DNA and monochromosomal controls. Enzymatic amplification of genes within the HBM region on 11q12-13 was accomplished using the PCR with oligonucleotides
20 flanking each exon as well as the putative 5' regulatory elements of each gene. The primers were chosen to amplify each exon as well as 15 or more base pairs within each intron on either side of the splice. All PCR primers were made as chimeras to facilitate dye primer sequencing. The M13-21F (5'- GTA A CGA CGG CCA GT -3') (SEQ ID NO:30) and -
28REV (5'- AAC AGC TAT GAC CAT G -3') (SEQ ID NO:31) primer binding sites were
25 built on to the 5' end of each forward and reverse PCR primer, respectively, during synthesis. 150 ng of genomic DNA was used in a 50 µl PCR with 2 U AmpliTaq, 500 nM primer and 125 µM dNTP. Buffer and cycling conditions were specific to each primer set. TaqStart antibody (Clontech) was used for hot start PCR to minimize primer dimer formation. 10% of the product was examined on an agarose gel. The appropriate samples were diluted 1:25 with
30 deionized water before sequencing.

Each PCR product was sequenced according to the standard Energy Transfer primer (Amersham) protocol. All reactions took place in 96 well trays. 4 separate reactions, one each for A, C, G and T were performed for each template. Each reaction included 2 μ l of the sequencing reaction mix and 3 μ l of diluted template. The plates were then heat sealed with foil tape and placed in a thermal cycler and cycled according to the manufacturer's recommendation. After cycling, the 4 reactions were pooled. 3 μ l of the pooled product was transferred to a new 96 well plate and 1 μ l of the manufacturer's loading dye was added to each well. All 96 well pipetting procedures occurred on a Hydra 96 pipetting station (Robbins Scientific, USA). 1 μ l of pooled material was directly loaded onto a 48 lane gel running on an ABI 377 DNA sequencer for a 10 hour, 2.4 kV run.

Polyphred (University of Washington) was used to assemble sequence sets for viewing with *Consed* (University of Washington). Sequences were assembled in groups representing all relevant family members and controls for a specified target region. This was done separately for each of the three stages. Forward and reverse reads were included for each individual along with reads from the monochromosomal templates and a color annotated reference sequence. *Polyphred* indicated potential polymorphic sites with a purple flag. Two readers independently viewed each assembly and assessed the validity of the purple-flagged sites.

A total of 23 exons present in the mature mRNA and several other portions of the primary transcript were evaluated for heterozygosity in the nuclear family of two HBM-affected and two unaffected individuals. 25 SNPs were identified, as shown in the table below.

TABLE 8

Single Nucleotide Polymorphisms in the *Zmax1* (*LRP5*) Gene and Environs

Exon Name	Location	Base Change
b200e21-h_Contig1_1.nt	69169 (309G)	C/A
b200e21-h_Contig4_12.nt	27402 (309G)	A/G
b200e21-h_Contig4_13.nt	27841 (309G)	T/C
b200e21-h_Contig4_16.nt	35600 (309G)	A/G
b200e21-h_Contig4_21.nt	45619 (309G)	G/A

	Exon Name	Location	Base Change
	b200e21-h_Contig4_22.nt-a	46018 (309G)	T/G
5	b200e21-h_Contig4_22.nt-b	46093 (309G)	T/G
	b200e21-h_Contig4_22.nt-c	46190 (309G)	A/G
	b200e21-h_Contig4_24.nt-a	50993 (309G)	T/C
10	b200e21-h_Contig4_24.nt-b	51124 (309G)	C/T
	b200e21-h_Contig4_25.nt	55461 (309G)	C/T
15	b200e21-h_Contig4_33.nt-a	63645 (309G)	C/A
	b200e21-h_Contig4_33.nt-b	63646 (309G)	A/C
	b200e21-h_Contig4_61.nt	24809 (309G)	T/G
20	b200e21-h_Contig4_62.nt	27837 (309G)	T/C
	b200e21-h_Contig4_63.nt-a	31485 (309G)	C/T
25	b200e21-h_Contig4_63.nt-b	31683 (309G)	A/G
	b200e21-h_Contig4_9.nt	24808 (309G)	T/G
	b527d12-h_Contig030g_1.nt-a	31340 (308G)	T/C
30	b527d12-h_Contig030g_1.nt-b	32538 (308G)	A/G
	b527d12-h_Contig080C_2.nt	13224 (308G)	A/G
35	b527d12-h_Contig087C_1.nt	21119 (308G)	C/A
	b527d12-h_Contig087C_4.nt	30497 (308G)	G/A
	b527d12-h_Contig088C_4.nt	24811 (309G)	A/C
40	b527d12-h_Contig089_1HP.nt	68280 (309G)	G/A

In addition to the polymorphisms presented in Table 8, two additional polymorphisms can also be present in SEQ ID NO:2. These is a change at position 2002 of SEQ ID NO:2. Either a guanine or an adenine can appear at this position. This polymorphism is silent and is not associated with any change in the amino acid sequence. The second change is at position 4059 of SEQ ID NO:2 corresponding in a cytosine (C) to thymine (T) change. This polymorphism results in a corresponding amino acid change from a valine (V) to an alanine (A). Other polymorphisms were found in the candidate gene exons and adjacent intron sequences as displayed in Tables 2 and 3. Any one or combination of the polymorphisms

listed in Tables 2, 3 or 8 or the two discussed above could also have a minor effect on bone mass when present in SEQ ID NO:2. These could also be in combination with any of the other mutations discussed in Section 3 and in the Examples below.

The present invention encompasses the nucleic acid sequences having the nucleic acid sequence of SEQ ID NO: 1 with any one or more of the above-identified point mutations.

Preferably, the present invention encompasses the nucleic acid of SEQ ID NO: 2 or is a mutation in SEQ ID NO:1 that produces a phenotype like that observed with the protein encoded by SEQ ID NO:2 (HBM phenotype). Specifically, a base-pair substitution changing G to T at position 582 in the coding sequence of *Zmax1* (the *HBM* gene) was identified as heterozygous in all HBM individuals, and not found in the unaffected individuals (i.e., b527d12-h_Contig087C_1.nt). Fig. 6 shows the order of the contigs in B527D12. The direction of transcription for the *HBM* gene is from left to right. The sequence of contig308G of B527D12 is the reverse complement of the coding region to the *HBM* gene. Therefore, the relative polymorphism in contig 308G shown in Table 8 as a base change substitution of C to A is the complement to the G to T substitution in the *HBM* gene. This mutation causes a substitution of glycine 171 with valine (G171V).

The HBM polymorphism was confirmed by examining the DNA sequence of different groups of individuals. In all members of the HBM pedigree (38 individuals), the HBM polymorphism was observed in the heterozygous form in affected (i.e., elevated bone mass) individuals only (N=18). In unaffected relatives (N=20) (BMDZ<2.0), the HBM polymorphism was never observed. To determine whether this polymorphism was ever observed in individuals outside of the HBM pedigree, 297 phenotyped individuals were characterized at the site of the *HBM* gene. None were heterozygous at the site of the HBM polymorphism. In an unphenotyped control group, none of 64 individuals were observed to be heterozygous at position 582. Taken together, these data prove that the polymorphism observed in the kindred displaying the high bone mass phenotype is strongly correlated with the G→T polymorphism at position 582 of *Zmax1* (LRP5). Furthermore, these results coupled with the ASO results described below, establish that the HBM polymorphism genetically segregates with the HBM phenotype, and that both the HBM polymorphism and

phenotype are rare in the general population.

13. Allele Specific Oligonucleotide (ASO) Analysis

The amplicon containing the HBM polymorphism was PCR amplified using primers specific for the exon of interest. The appropriate population of individuals was PCR amplified in 96 well microtiter plates as follows. PCR reactions (20 μ l) containing 1X Promega PCR buffer (Cat. # M1883 containing 1.5 mM $MgCl_2$), 100 mM dNTP, 200 nM PCR primers (1863F: 5'-CCAAGTTCTGAGAAGTCC-3' and 1864R: 5'-AATACCTGAAACCATACCTG-3'; SEQ ID NOS: 629 and 630 respectively), 1 U AmpliTaqTM, and 20 ng of genomic DNA were prepared and amplified under the following PCR conditions: 94°C, 1 minute, (94°C, 30 sec.; 58°C, 30 sec.; 72°C, 1 min X35 cycles), 72°C, 5 min, 4°C, hold. Loading dye was then added and 10 μ l of the products was electrophoresed on 1.5% agarose gels containing 1 μ g/ml ethidium bromide at 100-150 V for 5-10 minutes. Gels were treated 20 minutes in denaturing solution (1.5 M NaCl, 0.5 N NaOH), and rinsed briefly with water. Gels were then neutralized in 1 M Tris-HCl, pH 7.5, 1.5 M NaCl, for 20 minutes and rinsed with water. Gels were soaked in 10 X SSC for 20 minutes and blotted onto nylon transfer membrane (Hybond N+- Amersham) in 10X SSC overnight. Filters were the rinsed in 6X SSC for 10 minutes and UV crosslinked.

The allele specific oligonucleotides (ASO) were designed with the polymorphism approximately in the middle. Oligonucleotides were phosphate free at the 5' end and were purchased from Gibco BRL. Sequences of the oligonucleotides are:

2326 Zmax1.ASO.g: 5'-AGACTGGGGTGGAGACGC-3' (SEQ ID NO:631)

2327 Zmax1.ASO.t: 5'-CAGACTGGGGTTGAGACGCC-3' (SEQ ID NO:632)

The polymorphic nucleotides are underlined. To label the oligos, 1.5 μ l of 1 μ g/ μ l ASO oligo (2326.Zmax1.ASO.g or 2327.Zmax1.ASO.t), 11 μ l ddH₂O, 2 μ l 10X kinase forward buffer, 5 μ l γ^{32} P-ATP (6000 Ci/mMole), and 1 μ l T4 polynucleotide kinase (10 U/ μ l) were mixed, and the reaction incubated at 37°C for 30-60 minutes. Reactions were then placed at 95°C for 2 minutes and 30 ml H₂O was added. The probes were purified using a G25 microspin column (Pharmacia).

Blots were prehybridized in 10 ml 5X SSPE, 5X Denhardt's, 2% SDS, and 100 µg/ml, denatured, sonicated salmon sperm DNA at 40°C for 2 hr. The entire reaction mix of kinased oligo was then added to 10 ml fresh hybridization buffer (5X SSPE, 5X Denhardt's, 2% SDS) and hybridized at 40°C for at least 4 hours to overnight.

5 All washes done in 5X SSPE, 0.1 % SDS. The first wash was at 45°C for 15 minutes; the solution was then changed and the filters washed 50°C for 15 minutes. Filters were then exposed to Kodak biomax film with 2 intensifying screens at -70°C for 15 minutes to 1 hr. If necessary the filters were washed at 55°C for 15 minutes and exposed to film again. Filters were stripped by washing in boiling 0.1X SSC, 0.1% SDS for 10 minutes at least 3
10 times.

The two films that best captured the allele specific assay with the 2 ASOs were converted into digital images by scanning them into Adobe PhotoShop. These images were overlaid against each other in Graphic Converter and then scored and stored in FileMaker Pro 4.0 (see Fig. 9).

15 In order to determine the HBM allele frequency in ethnically diverse populations, 672 random individuals from various ethnic groups were typed by the allele specific oligonucleotide (ASO) method. This population included 96 CEPH grandparents (primarily Caucasian), 192 Caucasian, 192 African-American, 96 Hispanic, and 96 Asian individuals. No evidence was obtained for the presence of the HBM polymorphism in any of these
20 individuals. Overall, a total of 911 individuals were typed either by direct sequencing or ASO hybridization; all were homozygous GG at the site of the HBM polymorphism (Fig. 14). This information illustrates that the HBM allele is rare in various ethnic populations.

Thus this invention provides a rapid method of identifying individuals with the HBM allele. This method could be used in the area of diagnostics and screening of an individual
25 for susceptibility to osteoporosis or other bone disorder. The assay could also be used to identify additional individuals with the HBM allele or the additional polymorphisms described herein.

14. Cellular Localization of Zmax1 (LRP5)

14.1 Gene Expression in rat tibia by non isotopic *in situ* Hybridization

In situ hybridization was conducted by Pathology Associates International (PAI), Frederick, MD. This study was undertaken to determine the specific cell types that express the *Zmax1 (LRP5)* gene in rat bone with particular emphasis on areas of bone growth and remodeling. *Zmax1 (LRP5)* probes used in this study were generated from both human (HuZmax1) and mouse (MsZmax1) cDNAs, which share an 87% sequence identity. The homology of human and mouse *Zmax1 (LRP5)* with rat *Zmax1 (LRP5)* is unknown.

For example, gene expression by non-isotopic *in situ* hybridization was performed as follows, but other methods would be known to the skilled artisan. Tibias were collected from two 6 to 8 week old female Sprague Dawley rats euthanized by carbon dioxide asphyxiation. Distal ends were removed and proximal tibias were snap frozen in OCT embedding medium with liquid nitrogen immediately following death. Tissues were stored in a -80°C freezer.

Probes for amplifying PCR products from cDNA were prepared as follows. The primers to amplify PCR products from a cDNA clone were chosen using published sequences of both human LRP5 (GenBank Accession No. ABO17498) and mouse LRP5 (GenBank Accession No. AF064984). In order to minimize cross reactivity with other genes in the LDL receptor family, the PCR products were derived from an intracellular portion of the protein coding region. PCR was performed in a 50 μl reaction volume using cDNA clone as template. PCR reactions contained 1.5 mM MgCl_2 , 1 unit AmpliTaq, 200 μM dNTPs and 2 μM each primer. PCR cycling conditions were 94°C for 1 min., followed by 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds; followed by a 5 minute extension at 72°C . The reactions were then run on a 1.5% agarose Tris-Acetate gel. DNA was eluted from the agarose, ethanol precipitated and resuspended in 10 mM Tris, pH 8.0.

Gel purified PCR products were prepared for both mouse and human cDNAs and supplied to

Pathology Associates International for *in situ* hybridizations.

The sequence of the human and mouse PCR primers and products were as follows:

Human Zmax1 (LRP5) sense primer (HBM253)

5'-CCCGTGTGCTCCGCCGCCAGTTC-3' (SEQ ID NO:633)

5 Human Zmax1 (LRP5) antisense primer (HBM465)

5'-GGCTCACGGAGCTCATCATGGACTT-3' (SEQ ID NO:634)

Human Zmax1 PCR product

5'-CCCGTGTGCTCCGCCGCCAGTTCCTGCGCGCGGGGTCAGTGTGTGGACCT
GCGCCTGCGCTGCGACGGCGAGGCAGACTGTCAGGACCGCTCAGACGAGGTGG
10 ACTGTGACGCCATCTGCCTGCCCAACCAGTTCGGGTGTGCGAGCGGCCAGTGTGT
CCTCATCAAACAGCAGTGCGACTCCTTCCCCGACTGTATCGACGGCTCCGACGAG
CTCATGTGTGAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAGCAGT
GCCATCGGGCCCGTCATTGGCATCATCCTCTCTCTCTTCGTCATGGGTGGTGTCTA
TTTTGTGTGCCAGCGCGTGGTGTGCCAGCGCTATGCGGGGGCCAACGGGGCCCTTC
15 CCGCACGAGTATGTCAGCGGGACCCCGCACGTGCCCTCAATTCATAGCCCCGG
GCGGTTCCCAGCATGGCCCCCTCACAGGCATCGCATGCGGAAAGTCCATGATGA
GCTCCGTGAGCC-3' (SEQ ID NO:635)

Mouse Zmax1 (LRP5) Sense primer (HBM655)

5'-AGCGAGGCCACCATCCACAGG-3' (SEQ ID NO:636)

20 Mouse Zmax1 (LRP5) antisense primer (HBM656)

5'-TCGCTGGTCGGCATAATCAAT-3' (SEQ ID NO:637)

Mouse (LRP5) 1 PCR product

5'-AGCAGAGCCACCATCCACAGGATCTCCCTGGAGACTAACAACAACGATGTGG
CTATCCCACCTCACGGGTGTCAAAGAGGCCTCTGCACTGGACTTTGATGTGTCCAA

CAATCACATCTACTGGACTGATGTTAGCCTCAAGACGATCAGCCGAGCCTTCATG
 AATGGGAGCTCAGTGGAGCACGTGATTGAGTTTGGCCTCGACTACCCTGAAGGA
 ATGGCTGTGGACTGGATGGGCAAGAACCTCTATTGGGCGGACACAGGGACCAAC
 AGGATTGAGGTGGCCCGGCTGGATGGGCAGTTCCGGCAGGTGCTTGTGTGGAGA
 5 GACCTTGACAACCCCAAGGTCTCTGGCTCTGGATCCTACTAAAGGCTACATCTACT
 GGACTGAGTGGGGTGGCAAGCCAAGGATTGTGCGGGCCTTCATGGATGGGACCA
 ATTGTATGACACTGGTAGACAAGGTGGGCCGGGCCAACGACCTCACCATTGATT
 ATGCCGACCAGCGA-3' (SEQ ID NO:638)

Riboprobes were synthesized as follows. The PCR products were reamplified with
 10 chimeric primers designed to incorporate either a T3 promoter upstream, or a T7 promoter
 downstream of the reamplification products. The resulting PCR products were used as
 template to synthesize digoxigenin-labeled riboprobes by *in vitro* transcription (IVT).
 Antisense and sense riboprobes were synthesized using T7 and T3 RNA polymerases,
 respectively, in the presence of digoxigenin-11-UTP (Boehringer-Mannheim) using a
 15 MAXIscript IVT kit (Ambion) according to the manufacturer. The DNA was then degraded
 with Dnase-1, and unincorporated digoxigenin was removed by ultrafiltration. Riboprobe
 integrity was assessed by electrophoresis through a denaturing polyacrylamide gel.
 Molecular size was compared with the electrophoretic mobility of a 100–1000 base pair (bp)
 RNA ladder (Ambion). Probe yield and labeling was evaluated by blot immunochemistry.
 20 Riboprobes were stored in 5 µl aliquots at –80°C.

The *in situ* hybridization was performed as follows. Frozen rat bone was cut into 5
 µM sections on a Jung CM3000 cryostat (Leica) and mounted on adhesive slides
 (Instrumedics). Sections were kept in the cryostat at –20°C until all the slides were prepared
 in order to prevent mRNA degradation prior to post-fixation for 15 minutes in 4%
 25 paraformaldehyde. Following post-fixation, sections were incubated with 1 ng/µl of either

antisense or sense riboprobe in Pathology Associates International (PAI) customized hybridization buffer for approximately 40 hours at 58°C. Following hybridization, slides were subjected to a series of post-hybridization stringency washes to reduce nonspecific probe binding. Hybridization was visualized by immunohistochemistry with an anti-
5 digoxigenin antibody (FAB fragment) conjugated to alkaline phosphatase. Nitroblue tetrazolium chloride/bromochloroindolyl phosphate (Boehringer-Mannheim), a precipitating alkaline phosphatase substrate, was used as the chromogen to stain hybridizing cells purple to nearly black, depending on the degree of staining. Tissue sections were counter-stained with nuclear fast red. Assay controls included omission of the probe, omission of probe and anti-
10 digoxigenin antibody.

Specific cell types were assessed for demonstration of hybridization with antisense probes by visualizing a purple to black cytoplasmic and/or peri-nuclear staining indicating a positive hybridization signal for mRNA. Each cell type was compared to the replicate sections, which were hybridized with the respective sense probe. Results were considered
15 positive if staining was observed with the antisense probe and no staining or weak background with the sense probe.

The cellular localization of the hybridization signal for each of the study probes is summarized in Table 9. Hybridization for Zmax1 (LRP5) was primarily detected in areas of bone involved in remodeling, including the endosteum and trabecular bone within the
20 metaphysis. Hybridization in selected bone lining cells of the periosteum and epiphysis were also observed. Positive signal was also noted in chondrocytes within the growth plate, particularly in the proliferating chondrocytes. See Figs. 10, 11 and 12 for representative photomicrographs of *in situ* hybridization results.

TABLE 9

Summary of Zmax1 (LRP5) *in situ* hybridization in rat tibia

PROBE	SITE	ISH SIGNAL
Hu Zmax1	<u>Epiphysis</u>	
	Osteoblasts	+
	Osteoclasts	-
	<u>Growth Plate</u>	
	resting chondrocytes	-
	proliferating chondrocytes	+
	hypertrophic chondrocytes	-
	<u>Metaphysis</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Diaphysis</u>	-
	<u>Endosteum</u>	
	osteoblasts	+
	osteoclasts	+
	Periosteum	-
	<u>MsZmax1</u>	
	<u>Epiphysis</u>	
	Osteoblasts	+
	Osteoclasts	-
	<u>Growth Plate</u>	
	resting chondrocytes	-
	proliferating chondrocytes	+
	hypertrophic chondrocytes	+
	<u>Metaphysis</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Diaphysis</u>	-
	<u>Endosteum</u>	
	osteoblasts	+
	osteoclasts	+
	Periosteum	+

Legend: "+" = hybridization signal detected "-" = no hybridization signal detected

"ISH" – *In situ* hybridization

These studies confirm the positional expression of Zmax1 (LRP5) in cells involved in bone remodeling and bone formation. Zmax1 (LRP5) expression in the zone of proliferation

and in the osteoblasts and osteoclasts of the proximal metaphysis, suggests that the *Zmax1* (*LRP5*) gene is involved in the process of bone growth and mineralization. The activity and differentiation of osteoblasts and osteoclasts are closely coordinated during development as bone is formed and during growth as well as in adult life as bone undergoes continuous remodeling. The formation of internal bone structures and bone remodeling result from the coupling of bone resorption by activated osteoclasts with subsequent deposition of new material by osteoblasts. *Zmax1* (*LRP5*) is related to the LDL receptor gene, and thus may be a receptor involved in mechanosensation and subsequent signaling in the process of bone remodeling. Therefore, changes in the level of expression of this gene could impact on the rate of remodeling and degree of mineralization of bone.

15. **Antisense**

Antisense oligonucleotides are short synthetic nucleic acids that contain complementary base sequences to a targeted RNA. Hybridization of the RNA in living cells with the antisense oligonucleotide interferes with RNA function and ultimately blocks protein expression. Therefore, any gene for which the partial sequence is known can be targeted by an antisense oligonucleotide.

Antisense technology is becoming a widely used research tool and will play an increasingly important role in the validation and elucidation of therapeutic targets identified by genomic sequencing efforts.

Antisense technology was developed to inhibit gene expression by utilizing an oligonucleotide complementary to the mRNA that encodes the target gene. There are several possible mechanisms for the inhibitory effects of antisense oligonucleotides. Among them, degradation of mRNA by RNase H is considered to be the major mechanism of inhibition of protein function. This technique was originally used to elucidate the function of a target gene, but may also have therapeutic applications, provided it is designed carefully and properly.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed

using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, β -D-galactosylqueosine, inosine, N6-isopentenyladenine, I-methylguanine, I-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), t-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

In addition, the use of morpholino oligonucleotides could be employed. Morpholinos are oligomers with modification of the ribose moiety to a morpholino group. This technology is covered by U.S. Patent No. 5,185,444 and is described in Summerton *et al.*, 1997 *Antisense Nucleic Acid Drug Dev.* 7(3): 187-95. The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an HBM or Zmax1 (LRP5) protein or a protein which interacts with Zmax1 (LRP5) and/or HBM to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration

of an antisense nucleic acid molecule of the invention includes direct injection at a tissue site. Alternatively, an antisense nucleic acid molecule can be modified to target selected cells and then administered systemically. For example, for systemic administration, an antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An μ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual γ -units, the strands run parallel to each other (Gaultier *et al.*, 1987 *Nucl. Acids Res.* 15: 6625-41). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987 *Nucl. Acids Res.* 15: 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987 *FEBS Lett.* 215: 327-330). In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff *et al.*, 1988 *Nature* 334: 585-591) can be used to catalytically cleave Zmax1 (LRP5) or HBM mRNA transcripts to thereby inhibit translation of Zmax1 (LRP5) or HBM mRNA. A ribozyme having specificity for a Zmax1- or HBM-encoding nucleic acid can be designed based upon the nucleotide sequence of a Zmax1 (LRP5) or HBM cDNA disclosed herein (i.e., SEQ ID NOS:1 or 3). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an HBM or Zmax1-encoding mRNA. See, e.g., Cech *et al.* U.S. Patent No. 4,987,071 and Cech *et al.* U.S. Patent No. 5,116,742 both incorporated herein by reference. Alternatively, Zmax1 (LRP5) or HBM mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel *et al.*, 1993 *Science* 261: 1411-1418. Alternatively, Zmax1 (LRP5) or HBM gene expression can be inhibited by targeting nucleotide sequences

complementary to the regulatory region of the *Zmax1* (*LRP5*) or *HBM* gene (e.g., the *Zmax1* or *HBM* gene promoter and/or enhancers) to form triple helical structures that prevent transcription of the *Zmax1* (*LRP5*) or *HBM* genes in target cells. See generally, Helene, 1991 *Anticancer Drug Des.* 6(6):569-84; Helene *et al.*, 1992 *Ann. N.Y. Acad. Sci.* 660:27-36; and
5 Maher, 1992 *Bioassays* 14(12):807-15.

Zmax1 (*LRP5*), *LRP6*, *HBM*-like and *HBM* gene expression can also be inhibited using RNA interference (RNAi) caused by small inhibitory RNAs (siRNAs). This is a technique for post-transcriptional gene silencing (PTGS), in which target gene activity is specifically abolished with cognate double-stranded RNA (dsRNA). siRNAs resemble in
10 many aspects PTGS in plants and has been detected in many invertebrates including trypanosome, hydra, planaria, nematode and fruit fly (*Drosophila melanogaster*). It may be involved in the modulation of transposable element mobilization and antiviral state formation. RNAi in mammalian systems is disclosed in PCT application WO 00/63364 which is incorporated by reference herein in its entirety. See also, Elbashir *et al.*, 2001
15 *Nature* 411: 494-98. Basically, dsRNA, homologous to the target (*Zmax1* or *HBM*) is introduced into the cell and a sequence specific reduction in gene activity is observed.

Another embodiment contemplates the use of small hairpin RNAs (shRNAs). These compounds are described further in Yu *et al.*, 2002 *Proc. Natl. Acad. Sci. USA* 99: 6047-52; and Paddison *et al.*, 2002 *Genes & Devel.* 16: 948-58.

As an example, preparing antisense oligonucleotides can be performed as follows. Studies have been undertaken using antisense technology in the osteoblast-like murine cell line, MC3T3. These cells can be triggered to develop along the bone differentiation sequence. An initial proliferation period is characterized by minimal expression of differentiation markers and initial synthesis of collagenous extracellular matrix. Collagen
25 matrix synthesis is required for subsequent induction of differentiation markers. Once the matrix synthesis begins, osteoblast marker genes are activated in a clear temporal sequence: alkaline phosphatase is induced at early times while bone sialoprotein and osteocalcin appear later in the differentiation process. This temporal sequence of gene expression is useful in monitoring the maturation and mineralization process. Matrix mineralization, which does not

begin until several days after maturation has started, involves deposition of mineral on and within collagen fibrils deep within the matrix near the cell layer-culture plate interface. The collagen fibril-associated mineral formed by cultured osteoblasts resembles that found in woven bone *in vivo* and therefore is used frequently as a study reagent.

5 MC3T3 cells were transfected with antisense oligonucleotides for the first week of the differentiation, according to the manufacturer's specifications (U.S. Patent No. 5,849,902).

The oligonucleotides designed for Zmax1 (LRP5) are given below:

10875: 5'-AGUACAGCUUCUUGCCAACCCAGUC-3' (SEQ ID NO:639)

10876: 5'-UCCUCCAGGUCGAUGGUCAGCCCAU-3' (SEQ ID NO:640)

10 10877: 5'-GUCUGAGUCCGAGUUCAAAUCCAGG-3' (SEQ ID NO:641)

Fig. 13 shows the results of antisense inhibition of (LRP5) 1 in MC3T3 cells. The three oligonucleotides shown above were transfected into MC3T3 and RNA was isolated according to standard procedures. Northern analysis clearly shows markedly lower steady state levels of the Zmax1 (LRP5) transcript while the control gene GAPDH remained
15 unchanged. Thus, antisense technology using the primers described above allows for the study of the role of Zmax1 (LRP5) expression on bone biology.

16. Yeast Two Hybrid

In order to identify the signaling pathway that Zmax1 (LRP5) participates in to
20 modulate bone density, the yeast two hybrid protein interaction technology was utilized. This technique facilitates the identification of proteins that interact with one another by coupling tester proteins to components of a yeast transcription system (Fields *et al.*, 1989 *Nature* 340: 245-246; U.S. Pat. No. 5,283,173 by Fields *et al.*; Johnston, 1987 *Microbiol. Rev.* 51: 458-476; Keegan *et al.*, 1986 *Science* 231: 699-704; Durfee *et al.*, 1993, *Genes Dev.*
25 7: 555-569; Chien *et al.*, 1991 *Proc. Natl. Acad. Sci USA* 88: 9578-9582; Fields *et al.*, 1994 *Trends in Genetics* 10: 286-292; and Gyuris *et al.*, 1993 *Cell* 75: 791-803). First a "bait" protein, the protein for which one seeks interacting proteins, is fused to the DNA binding domain of a yeast transcription factor. Second, a cDNA library is constructed that contains cDNAs fused to the transcriptional activation domain of the same yeast transcription factor;

this is termed the prey library. The bait construct and prey library are transformed into yeast cells and then mated to produce diploid cells. If the bait interacts with a specific prey from the cDNA library, the activation domain is brought into the vicinity of the promoter via this interaction. Transcription is then driven through selectable marker genes and growth on selective media indicates the presence of interacting proteins.

The amino acid sequence used in the yeast two hybrid experiments discussed herein consisted of the entire cytoplasmic domain and a portion of the transmembrane domain and is shown below (amino to carboxy orientation):

RVVCQRYAGA NGPFPHEYVS GTPHVPLNFI APGGSQHGPF TGIACGKSMM
SSVSLMGGRG GVPLYDRNHV TGASSSSSSS TKATLYPPIL NPPSPATDP
SLYNMDMFYS SNIPATVRPY RPYYIIRGMAP PTTPCSTDVC DSDYSASRWK
ASKYYLDLNS DSDPYPPPPT PHSQYLSAED SCPPSPATER SYFHLFPPPP
SPCTDSS (SEQ ID NO: 765)

The last 6 amino acids of the putative transmembrane domain are indicated in bold. Putative SH3 domains are underlined. Additional amino acid sequences of 50 amino acids or greater in either the proteins encoded by the Zmax1 (LRP5) or HBM alleles can also be used as bait. The upper size of the polypeptide used as bait is limited only by the presence of a complete transmembrane domain (see Fig. 4), which will render the bait to be nonfunctional in a yeast two hybrid system. These additional bait proteins can be used to identify additional proteins which interact with the proteins encoded by HBM or Zmax1 (LRP5) in the focal adhesion signaling pathway or in other pathways in which these HBM or Zmax1 (LRP5) proteins may act. Once identified, methods of identifying agents which regulate the proteins in the focal adhesion signaling pathway or other pathways in which HBM acts can be performed as described herein for the HBM and Zmax1 (LRP5) proteins.

In order to identify cytoplasmic Zmax1 (LRP5) signaling pathways, the cytoplasmic domain of Zmax1 (LRP5) was subcloned into two bait vectors. The first vector was pDBleu, which was used to screen a brain, and Hela prey cDNA library cloned into the vector pPC86 (Clontech). The second bait vector used was pDBtrp, which was used to screen a cDNA prey library derived from the TE85 osteosarcoma cell line in vector pOP46. Another suitable vector which is widely available, is p86 (Gibco, iestTM System). Standard techniques known

to those skilled in the art were used as described in Fields and Song, 1989 *Nature* 340: 245-246; U.S. Pat. No. 5,283,173 by Fields and Song; Johnston, 1987 *Microbiol. Rev.* 51: 458-476; Keegan *et al.*, 1986 *Science* 231: 699-704; Durfee *et al.*, 1993 *Genes Dev.* 7: 555-569; Chien *et al.*, 1991 *Proc. Natl. Acad. Sci USA* 88: 9578-9582; Fields *et al.*, 1994 *Trends in Genetics* 10: 286-292; and Gyuris *et al.*, 1993 *Cell* 75: 791-803. The bait construct and prey cDNA libraries were transformed into yeast using standard procedures.

To perform the protein interaction screen, an overnight culture of the bait yeast strain was grown in 20 ml SD selective medium with 2% glucose (pDBLeu, SD -Leu medium, pDBtrp, SD -trp medium). The cultures were shaken vigorously at 30°C overnight. The cultures were diluted 1:10 with complete medium (YEPD with 2% glucose) and the cultures then incubated with shaking for 2 hrs at 30°C.

The frozen prey library was thawed, and the yeast cells reactivated by growing them in 150 ml YEPD medium with 2% glucose for 2 hrs at 30°C. A filter unit was sterilized with 70% ethanol and washed with sterile water to remove the ethanol. The cell densities of both bait and prey cultures were measured by determining the OD at 600 nm. An appropriate volume of yeast cells that corresponded to a cell number of 1 ml of OD 600 = 4 of each yeast strain, bait and prey (library) was placed in a 50 ml Falcon tube. The mixture was then filtered through the sterilized filter unit. The filter was then transferred onto a prewarmed YEPD agar plate with the cell side up, removing all air bubbles underneath the filter. Plates were then incubated at 30°C for 6 hrs. One filter was transferred into a 50 ml Falcon tube, and 10 ml of SD with 2% Glucose was added; cells were resuspended by vortexing for 10 sec.

The number of primary diploid cells (growth on SD -Leu, -Trp plates) versus the numbers of colony forming units growing on SD -Trp and SD -Leu plates only was then titered. Different dilutions were plated and incubated at 30°C for two days. The number of colony forming units was then counted. The number of diploid colonies (colonies on SD -Leu -Trp plates) permits the calculation of whether or not the whole library of prey constructs was mated to the yeast expressing the bait. This information is important to judge the quality of the screen.

16.1 Indirect selection

Resuspended cells from 5 filter-matings were then pooled and the cells sedimented by centrifugation in a 50 ml Falcon tube. Cells were then resuspended in 16 ml SD medium with 2% Glc. Two ml of this cell suspension was plated onto 8 square plates each (SD -Leu, -
5 Trp) with sterile glass beads and selected for diploid cells by incubating at 30°C for 18 - 20 hrs.

Cells were then scraped off the square plates, the cells centrifuged and combined into one 50 ml Falcon tube. The cell pellet was then resuspended in 25 ml of SD medium with 2% glucose. The cell number was then determined by counting of an appropriate dilution
10 (usually 1:100 to 1:1000) with a Neugebauer chamber. Approximately 5×10^7 diploid cells were plated onto the selective medium. The observations about the growth of the bait strain together with irrelevant prey vectors helps to determine which selective plates will have to be chosen for the library screen. Generally, all screens were plated on one square plate each with SD -Leu, -Trp, -His; SD -Leu, -Trp, His, 5 mM 3AT, and SD -Leu, -Trp, -His, -Ade is
15 recommended.

The yeast cells were then spread homogeneously with sterile glass beads and incubated at 30°C for 4 days. The number of colony forming yeast cells was titered by plating different dilutions of the scraped cell suspension onto SD -Leu, -Trp plates. Usually, plating of 100 μ l of a 10^{-3} and 10^{-4} dilution gave 100 - 1000 colonies per plate.
20

16.2 Direct selection

Five filters with the mated yeast cells were each transferred into separate 50 ml Falcon tubes and the cells resuspended with 10 ml SD medium with 2% Glc, each, followed by vortexing for 10 sec. The resuspended cells were combined and centrifuged in a Beckman
25 centrifuge at 3000 rpm. The supernatant was discarded and the cells resuspended in 6 ml of SD medium with 2% Glc. Two ml of the suspension was spread onto each selective square plate and incubated at 30°C for 4 - 5 days.

16.3 Isolation of Single Colonies

Yeast cells from an isolated colony were picked with a sterile tooth pick and transferred into individual wells of a 96 well plate. The cells were resuspended in 50 μ l of SD -Leu, -Trp, -His medium and incubated at 30°C for one day. The yeast cells were then stamped onto a SD -Leu, -Trp, -His plate in 96 well format and incubated at 30°C for 2 days. Yeast cells were also stamped onto a Nylon filter covering a YEPD plate and incubated at 30°C for one day. The cells on the Nylon filter were used for the analysis of the β - Gal reporter activity.

Yeast colonies were scraped from the SD -Leu, -Trp, -His plate with a sterile tooth pick, and reconfigured, if necessary, according to the β -Gal activity and then resuspended in 20 % glycerol. This served as a master plate for storage at -80°C.

For DNA preparation, yeast cells from the glycerol stock were stamped onto a SD -Trp plate and incubated at 30°C for 2 days. After two days of incubation, the yeast colonies were ready for colony PCR and sequencing. Standard colony PCR conditions were used to amplify inserts from preys recovered from the interaction screen. Sequencing was done using standard sequencing reactions and ABI377 (Perkin Elmer) fluorescent sequencing machines.

16.4 Verification of bait/prey interaction

Glycerol stocks of the prey of interest were thawed and inoculated in a 10 ml overnight culture of SD with glucose -Trp. After overnight growth, plasmid DNA preparation was performed using the BIO 101 RPM Yeast Plasmid Isolation Kit with 10 ml of culture. The culture was centrifuged and transferred to a 1.5 ml microcentrifuge tube. Yeast Lysis Matrix was then added to the pellet followed by 250 μ l of Alkaline Lysis Solution. Samples were then vortexed for 5 minutes. 250 μ l Neutralizing Solution was added and the sample mixed briefly. Samples were centrifuged for 2 minutes at room temperature in a microcentrifuge. The supernatant was transferred to a Spin Filter avoiding debris and Lysis Matrix. 250 μ l of Glassmilk Spin Buffer was added, and the tubes inverted to mix. Samples were centrifuged for 1 min and the liquid in the Catch Tube was discarded.

500 μ l of Wash Solution was added, the samples were centrifuged for 1 min, and the wash solution was discarded. The wash step was repeated once followed by a 1 min dry centrifugation to drive the remaining liquid out of the Spin Filter. The filter was transferred to a new Catch Tube and 100 μ l of sterile H₂O was added; samples were then vortexed briefly to resuspend and centrifuged for 30 seconds to collect the DNA in the bottom of the Catch Tube.

Five μ l of DNA was then transformed into DH10B Electromax cells using standard procedures and glycerol stocks prepared. Miniprep DNA was prepared using the Qiagen QIAprep Spin Miniprep Kit. DNA was finally eluted with 30 μ l of Qiagen EB buffer. One μ l of the plasmid DNA samples was then used to transform yeast cells using standard procedures. After 2 days of growth on SD –trp media, colonies were picked and patched onto fresh media. Similarly, bait colonies were patched onto SD –Leu media. Both were grown overnight at 30°C.

For mating, cells from bait and prey patches were spread together on YAPD media and incubated at 30°C for 12 hr. This plate was then replica plated onto an SD Agar-Leu-Trp plate and grown for 2 days at 30°C. To test the strength of interaction these plates were replica plated onto SD Agar-Leu-Trp-His, SD Agar-Leu-Trp-His with 5 mM 3AT and 10 mM 3AT, SD Agar-Leu-Trp-His-Ade, and SD Agar-Leu-Trp-Ura media and grown for 2 days at 30°C.

16.5 Galacton Star β -Galactosidase Activity Assay

After streaking and replica plating positive interactors on selection plates, colonies were placed in a 96 well dish with 200 μ l of SD-medium, leaving wells 1 and 96 blank. Ten microliters from the first 96 well dish was plated into another flat bottom 96 well dish containing 100 μ l of SD-medium. Controls consisted of a negative control and a very weak positive control. The cell density was measured at OD₆₀₀ (a value of 1 corresponds to 1×10^7 cells utilizing a 96 well spectrophotometer). The OD was usually between 0.03 and 0.10. Using microplates specifically for the luminometer, 50 μ l of reaction mixture were pipetted into each well. Fifty microliters of culture were then added and mixed by pipetting up and

down twice. The reaction was incubated for 30 minutes at room temperature followed by measurement of Relative Light Units using a luminometer.

Table 9 lists the genes identified in the yeast two hybrid screens from the 3 prey libraries tested. Two genes, zyxin and axin, were found to interact with the cytoplasmic domain of (LRP5) 1 in all three screens. Three genes, alpha-actinin, TCB and S1-5 interacted in two of the three screens.

A variety of proteins found at sites of cell-cell and cell-matrix contact (focal contacts/adhesion plaques) were shown to interact with the cytoplasmic domain of Zmax1 (LRP5). These include α -actinin, Trio, Pinch-like protein, and Zyxin. PINCH is a LIM domain-containing protein that is known to interact with integrin-linked kinase, an early signaler in integrin and growth factor signaling pathways. The finding of a closely related gene in the yeast two hybrid screen raises the possibility of a novel pathway linked to integrin signaling from extracellular matrix signals. Trio, also known to localize to focal adhesions, is thought to play a key role in coordinating cell-matrix interactions and cytoskeletal rearrangements involved in cell movement. Zyxin, another LIM domain-containing protein, is also localized to adhesion plaques and is thought to be involved in reorganization of the cytoskeleton when triggers are transmitted via integrin signaling pathways. Zyxin also interacts with alpha actinin, which we identified as interacting with Zmax1 (LRP5). Other LIM domain containing proteins identified include the human homologue of mouse ajuba, LIMD1, and a novel LIMD1-like protein.

Axin was also identified from the two hybrid experiments. This protein is involved in inhibition of the Wnt signaling pathway and interacts with the tumor suppressor APC. There is a link here with the focal adhesion signaling described above: one common step in the two pathways involves inhibition of glycogen synthase kinase 3, which in turn results in the activation of β -catenin/Lef-1 and AP-1 transcription factors. Axin/APC are involved in this as well as integrin linked kinase. The Wnt pathway has a role in determining cell fates during embryogenesis. If inappropriately activated, the Wnt pathway may also lead to cancer. The Wnt pathway also seems to have a role in cytoskeletal rearrangements. In a *Xenopus* embryo assay, the combination of HBM and Wnt5a proteins stimulated the Wnt pathway to a much

greater extent than the combination of Zmax1 (LRP5) and Wnt5a, which was modestly above the control and Wnt5a alone scores. The HBM and Zmax1 (LRP5) extracellular domains (ECD) caused a modest stimulation of Wnt signaling in the absence of Wnt5a which was slightly increased by the presence of Wnt5a in the presence of HBM ECD. A model depicting Zmax1 (LRP5) involvement in focal adhesion signaling is depicted in Fig. 15.

This data coupled with other studies suggest that integrin signaling pathways have a role in cellular responses to mechanical stress and adhesion. This provides an attractive model for the mechanism of action of Zmax1 (LRP5) in bone biology. It is possible that Zmax1 (LRP5) is involved in sensing either mechanical stress directly or binding a molecule in the extracellular matrix that is related to mechanical sensation. Signaling through subsequent pathways may be involved in bone remodeling due to effects on cell morphology, cell adhesion, migration, proliferation, differentiation, and apoptosis in bone cells.

TABLE 10: Yeast Two Hybrid Results

Gene Symbol	Gene	Genbank Accession #	NT SEQ ID NO:	AA SEQ ID NO:
ACTN1	alpha-actinin	NM_001102	642	
AES	amino-terminal enhancer of	NM_001130.3	643	
AIP4	atrophin-1 interacting protein	AF038564.1	644	
Novel	Ajuba		645	
AXIN	Wnt signaling	AF009674.1	646	
CDC23	cell division cycle 23, yeast, homolog	NM_004661.1	647	
HSM800944	Similar to TRIO	AL117435.1	748	
HSM800936		AL117427.1	649	
Novel	Similar to LIM domains containing protein 1.		650	
DEEPEST	mitotic spindle coiled-coil related protein	NM_006461.1	651	
ECM1	extracellular matrix protein 1	U65932.1	652	
EF1A	elongation factor 1-alpha	X16869.1	653	
FN	fibronectin	X02761.1	654	

5

10

15

20

25

Gene Symbol	Gene	Genbank Accession #	NT SEQ ID NO:	AA SEQ ID NO:
HOXB13	homeodomain protein	U81599.1	655	
Novel	Glu-Lys Rich protein		656	
LIMD1	LIM domains containing 1	NM_014240.1	657	
Novel	PINCH-like		658	
RANBPM	centrosomal protein	NM_005493.1	659	
S1-5	extracellular protein	U03877.1	660	
TCB	gene encoding cytosolic thyroid hormone-binding	M26252.1	661	
TID	tumorous imaginal discs	NM_005147.1	662	
ZYX	Zyxin	NM_003461.1	663	
TRIO	GTPase	U42390.1	664	
HUMPITPB	phosphatidylinositol transfer protein	D30037.1	665	
ACTN1	alpha-actinin	NP_001093.1		666
AES	amino-terminal enhancer of	NP_001121.2		667
AIP4	atrophin-1 interacting protein	AAC04845.1		668
Novel	Ajuba			669
AXIN	Wnt signaling	AAC51624.1		670
CDC23	cell division cycle 23, yeast homolog	NP_004652.1		671
Novel	Similar to TRIO CAB55923.1			672
Novel	Similar to LIM domains containing protein 1			673
DEEPEST	mitotic spindle coiled-coil related protein	NP_006452.1		674
ECM1	extracellular matrix protein 1	AAB05933.1		675
EF1A	elongation factor 1-alpha	CAA34756.1		676
FN	fibronectin	CAA26536.1		677
Novel	Glu-Lys rich protein			678
HOXB13	homeodomain protein B13	AAB39863.1		679
LIMD1	LIM domains containing 1	NP_055055.1		680
Novel	PINCH-like			681
RANBPM	centrosomal protein	NP_005484.1		682
S1-5	extracellular protein	AAA65590.1		683

Gene Symbol	Gene	Genbank Accession #	NT SEQ ID NO:	AA SEQ ID NO:
TCB	cytosolic thyroid hormone-binding protein	AAA36672.1		684
TID	tumorous imaginal discs	NP_005138.1		685
ZYX	Zyxin	NP_003452.1		686
TRIO	GTPase	AAC34245.1		687
PTDINSTP	phosphatidylinositol transfer protein beta isoform	P48739		688

In light of the model depicted in Fig. 15 and the results shown in Table 10, another aspect contemplated by the invention would be to regulate bone density and bone mass disorders by the regulating focal adhesion signaling. The regulation can occur by regulating the DNA, mRNA transcript or protein encoded by any of the members involved in the focal adhesion signaling pathway as identified by the yeast two hybrid system.

Also contemplated are the novel nucleic acids and proteins identified by the HBM yeast two hybrid system. These include but are not limited to SEQ ID NO: 645 (Ajuba), SEQ ID NO: 651 (a gene similar to a gene encoding LIM domains containing protein 1), SEQ ID NO: 656 (Glu-Lys Rich protein), SEQ ID NO: 658 (PINCH-like gene), SEQ ID NO: 669 (Ajuba protein), SEQ ID NO: 672 (protein similar to TRIO), SEQ ID NO: 673, SEQ ID NO: 678 (Glu-Lys rich protein) and SEQ ID NO: 681 (PINCH-like protein).

17. Potential Function

The protein encoded by Zmax1 (LRP5) and LRP6 are related to the Low Density Lipoprotein receptor (LDL receptor). *See, Goldstein et al., 1985 Ann. Rev. Cell Biology, 1: 1-39; Brown et al., 1986 Science, 232:34-47.* The LDL receptor is responsible for uptake of low density lipoprotein, a lipid-protein aggregate that includes cholesterol. Individuals with a defect in the LDL receptor are deficient in cholesterol removal and tend to develop atherosclerosis. In addition, cells with a defective LDL receptor show increased production of cholesterol, in part because of altered feedback regulation of cholesterol synthetic enzymes

and in part because of increased transcription of the genes for these enzymes. In some cell types, cholesterol is a precursor for the formation of steroid hormones.

Thus, the LDL receptor may, directly or indirectly, function as a signal transduction protein and may regulate gene expression. Because Zmax1 (LRP5) and LRP6 are related to the LDL receptor, this protein may also be involved in signaling between cells in a way that affects bone remodeling.

The glycine 171 amino acid is likely to be important for the function of Zmax1 (LRP5) because this amino acid is also found in the mouse homologue of Zmax1 (LRP5). The closely related LRP6 (Genbank Accession No. JE0272) protein also contains glycine at the corresponding position (Brown *et al.*, 1988 *Biochem. Biophys. Res. Comm.* 248: 879-88). Amino acids that are important in a protein's structure or function tend to be conserved between species, because natural selection prevents mutations with altered amino acids at important positions from arising.

In addition, the extracellular domain of Zmax1 (LRP5) contains four repeats consisting of five YWTD motifs followed by an EFG motif. This 5YWTD+EGF repeat is likely to form a distinct folded protein domain, as this repeat is also found in the LDL receptor and other LDL receptor-related proteins. The first three 5YWTD+EGF repeats are very similar in their structure, while the fourth is highly divergent. Glycine 171 occurs in the central YWTD motif of the first 5YWTD+EGF repeat in Zmax1 (LRP5). The other two similar 5YWTD+EGF repeats of Zmax1 (LRP5) also contain glycine at the corresponding position, as does the 5YWTD+EGF repeat in the LDL receptor protein. However, only 17.6% of the amino acids are identical among the first three 5YWTD+EGF repeats in Zmax1 (LRP5) and the single repeat in the LDL receptor. These observations indicate that glycine 171 is essential to the function of this repeat, and mutation of glycine 171 causes a functional alteration of Zmax1 (LRP5). The cDNA and peptide sequences are shown in Figs. 6A-6J. The critical base at nucleotide position 582 is indicated in bold and is underlined.

Northern blot analysis (Figs. 7A-B) reveals that Zmax1 (LRP5) is expressed in human bone tissue as well as numerous other tissues. A multiple-tissue Northern blot (Clontech, Palo Alto, CA) was probed with exons from Zmax1 (LRP5). As shown in Fig. 7A, the 5.5 kb

Zmax1 (LRP5) transcript was highly expressed in heart, kidney, lung, liver and pancreas and is expressed at lower levels in skeletal muscle and brain. A second northern blot, shown in Fig. 7B, confirmed the transcript size at 5.5 kb, and indicated that Zmax1 (LRP5) is expressed in bone, bone marrow, calvaria and human osteoblastic cell lines.

5 Taken together, these results coupled with the yeast two hybrid results indicate that the HBM polymorphism in the *Zmax1* (*LRP5*) gene is responsible for the HBM phenotype, and that the *Zmax1* gene is important in bone development. In addition, because mutation of *Zmax1* can alter bone mineralization and development, it is likely that molecules that bind to Zmax1 may usefully alter bone development. Such molecules may include, for example,
10 small molecules, proteins, RNA aptamers, peptide aptamers, and the like.

18. Preparation of Nucleic Acids, Vectors, Transformations and Host Cells

Large amounts of the nucleic acids of the present invention may be produced by replication in a suitable host cell. Natural or synthetic nucleic acid fragments coding for a
15 desired fragment will be incorporated into recombinant nucleic acid constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the nucleic acid constructs will be suitable for replication in a unicellular host, such as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cell lines.
20 The purification of nucleic acids produced by the methods of the present invention is described, for example, in Sambrook *et al.*, *Molecular Cloning. A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) or Ausubel *et al.*, *Current Protocols in Molecular Biology*, J. Wiley and Sons, NY (1992).

The nucleic acids of the present invention may also be produced by chemical
25 synthesis, e.g., by the phosphoramidite method described by Beaucage *et al.*, 1981 *Tetra. Lett.* 22: 1859-62 or the triester method according to Matteucci, *et al.*, 1981 *J. Am. Chem. Soc.* 103: 3185, and may be performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strands together

under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

Nucleic acid constructs prepared for introduction into a prokaryotic or eukaryotic host may comprise a replication system recognized by the host, including the intended nucleic acid fragment encoding the desired protein, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the protein encoding segment. Expression vectors may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Secretion signals may also be included where appropriate, whether from a native HBM or Zmax1 (LRP5) protein or from other receptors or from secreted proteins of the same or related species, which allow the protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Such vectors may be prepared by means of standard recombinant techniques well known in the art and discussed, for example, in Sambrook *et al.*, (1989) or Ausubel *et al.*, (1992).

An appropriate promoter and other necessary vector sequences will be selected so as to be functional in the host, and may include, when appropriate, those naturally associated with Zmax1 (LRP5) or HBM genes. Examples of workable combinations of cell lines and expression vectors are described in Sambrook *et al.*, (1989) or Ausubel *et al.*, (1992). Many useful vectors are known in the art and may be obtained from such vendors as Stratagene, New England BioLabs, Promega Biotech, and others. Promoters such as the trp, lac and phage promoters, tRNA promoters and glycolytic enzyme promoters may be used in prokaryotic hosts. Useful yeast promoters include promoter regions for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes such as enolase or glyceraldehyde-3-phosphate dehydrogenase, enzymes responsible for maltose and galactose utilization, and others. Vectors and promoters suitable for use in yeast expression are further described in EP 73,675A. Appropriate non-native mammalian promoters might include the early and late promoters from SV40 (Fiers *et al.*, 1978 *Nature* 273: 113) or promoters derived from murine

Moloney leukemia virus, mouse tumor virus, avian sarcoma viruses, adenovirus II, bovine papilloma virus or polyoma. In addition, the construct may be joined to an amplifiable gene (e.g., DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also *Enhancers and Eukaryotic Gene Expression*,
5 Cold Spring Harbor Press, Cold Spring Harbor, NY (1983).

While such expression vectors may replicate autonomously, they may also replicate by being inserted into the genome of the host cell, by methods well known in the art.

Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector.

10 The presence of this gene ensures growth of only those host cells which express the inserts. Typical selection genes encode proteins that a) confer resistance to antibiotics or other toxic substances, e.g. ampicillin, neomycin, methotrexate, etc.; b) complement auxotrophic deficiencies, or c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for Bacilli. The choice of the proper selectable marker will
15 depend on the host cell, and appropriate markers for different hosts are well known in the art.

The vectors containing the nucleic acids of interest can be transcribed *in vitro*, and the resulting RNA introduced into the host cell by well-known methods, e.g., by injection (see, Kubo *et al.*, *FEBS Letts.* 241: 119 (1988)), or the vectors can be introduced directly into host cells by methods well known in the art, which vary depending on the type of cellular host,
20 including electroporation; transfection employing calcium chloride, rubidium chloride, calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection; infection (where the vector is an infectious agent, such as a retroviral genome); and other methods. See generally, Sambrook *et al.*, 1989 and Ausubel *et al.*, 1992. The introduction of the nucleic acids into the host cell by any method known in the art, including
25 those described above, will be referred to herein as "transformation." The cells into which have been introduced nucleic acids described above are meant to also include the progeny of such cells.

Large quantities of the nucleic acids and proteins of the present invention may be prepared by expressing the Zmax1(LRP5), LRP6, HBM or HBM-like nucleic acids or

portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* or *Pseudomonas* may also be used.

Mammalian or other eukaryotic host cells, such as those of yeast, filamentous fungi,
5 plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. Propagation of mammalian cells in culture is per se well known. See, Jakoby and Pastan (eds.), *Cell Culture. Methods in Enzymology*, volume 58, Academic Press, Inc., Harcourt Brace Jovanovich, NY, (1979). Examples of commonly used mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cells,
10 and WI38, BHK, and COS cell lines, although it will be appreciated by the skilled practitioner that other cell lines may be appropriate, e.g., to provide higher expression desirable glycosylation patterns, or other features.

Clones are selected by using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the
15 same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of a particular product based on temperature sensitivity may also serve as an appropriate marker.

Prokaryotic or eukaryotic cells transformed with the nucleic acids of the present invention will be useful not only for the production of the nucleic acids and proteins of the
20 present invention, but also, for example, in studying the characteristics of Zmax1 (LRP5), LRP6, HBM or HBM-like proteins.

Antisense nucleic acid sequences are useful in preventing or diminishing the expression of Zmax1(LRP5), LRP6, HBM or HBM-like molecules, as will be appreciated by one skilled in the art. For example, nucleic acid vectors containing all or a portion of the
25 Zmax1(LRP5), LRP6, HBM or HBM-like nucleic acid may be placed under the control of a promoter in an antisense orientation and introduced into a cell. Expression of such an antisense construct within a cell will interfere with Zmax1(LRP5), LRP6, HBM or HBM-like transcription and/or translation and/or replication.

The probes and primers based on the *Zmax1(LRP5)*, *LRP6*, *HBM* or *HBM-like* gene

sequences disclosed herein are used to identify homologous gene sequences and proteins in other species. The gene sequences and proteins can also be used in the diagnostic/prognostic, therapeutic and drug screening methods described herein for the species from which they have been isolated.

5

19. Protein Expression and Purification

Expression and purification of the Zmax1 (LRP5), LRP6, HBM or HBM-like proteins of the invention can be performed essentially as outlined below (LRP5, LRP6 and HBM-like proteins are also included when referring only to HBM). To facilitate the cloning, expression and purification of membrane and secreted protein from the *HBM* gene, a gene expression system, such as the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli* was selected. Also, a DNA sequence encoding a peptide tag, the His-Tap, was fused to the 3' end of DNA sequences of interest to facilitate purification of the recombinant protein products. The 3' end was selected for fusion to avoid alteration of any 5' terminal signal sequence.

15

Nucleic acids chosen, for example, from the nucleic acids set forth in SEQ ID NOS: 1, 3 and 5-12 for cloning HBM were prepared by polymerase chain reaction (PCR). Synthetic oligonucleotide primers specific for the 5' and 3' ends of the HBM nucleotide sequence were designed and purchased from Life Technologies (Gaithersburg, MD). All forward primers (specific for the 5' end of the sequence) were designed to include an *NcoI* cloning site at the 5' terminus. These primers were designed to permit initiation of protein translation at the methionine residue encoded within the *NcoI* site followed by a valine residue and the protein encoded by the HBM DNA sequence. All reverse primers (specific for the 3' end of the sequence) included an *EcoRI* site at the 5' terminus to permit cloning of the HBM sequence into the reading frame of the pET-28b. The pET-28b vector provided a sequence encoding an additional 20 carboxyl-terminal amino acids including six histidine residues (at the C-terminus), which comprised the histidine affinity tag.

20

25

Genomic DNA prepared from the *HBM* gene was used as the source of template DNA for PCR amplification (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley

& Sons (1994)). To amplify a DNA sequence containing the HBM nucleotide sequence, genomic DNA (50 ng) was introduced into a reaction vial containing 2 mM MgCl₂, 1 μM synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined HBM, 0.2 mM of each of deoxynucleotide triphosphate, dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (AmpliTaq™, Roche Molecular Systems, Inc., Branchburg, NJ) in a final volume of 100 microliters.

Upon completion of thermal cycling reactions, each sample of amplified DNA was purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD). All amplified DNA samples were subjected to digestion with the restriction endonucleases, e.g., *Nco*I and *Eco*RI (New England BioLabs, Beverly, MA) (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)). DNA samples were then subjected to electrophoresis on 1.0% NuSeive (FMC BioProducts, Rockland, ME) agarose gels. DNA was visualized by exposure to ethidium bromide and long wave UV irradiation. DNA contained in slices isolated from the agarose gel was purified using the Bio 101 GeneClean Kit protocol (Bio 101, Vista, CA).

The pET-28b vector was prepared for cloning by digestion with restriction endonucleases, e.g., *Nco*I and *Eco*RI (New England BioLabs, Beverly, MA) (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)). The pET-28a vector, which encodes the histidine affinity tag that can be fused to the 5' end of an inserted gene, was prepared by digestion with appropriate restriction endonucleases.

Following digestion, DNA inserts were cloned (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)) into the previously digested pET-28b expression vector. Products of the ligation reaction were then used to transform the BL21 strain of *E. coli* (Ausubel *et al.*, 1994) as described below.

Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21 (DE3), were transformed with recombinant pET expression plasmids carrying the cloned HBM sequence according to standard methods (Ausubel *et al.*, 1994). Briefly, 1 μl of ligation reaction was mixed with 50 μl of electrocompetent cells and subjected to a high voltage pulse, after which samples were incubated in 0.45 ml SOC medium (0.5% yeast extract, 2.0% tryptone, 10 mM

NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 µg/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 were then picked and analyzed to evaluate cloned inserts, as described below.

5 Individual BL21 clones transformed with recombinant pET-28b HBM nucleotide sequences were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers specific for the HBM sequences that were used in the original PCR amplification cloning reactions. Successful amplification verifies the integration of the HBM sequence in the expression vector (Ausubel *et al.*, 1994).

10 Individual clones of recombinant pET-28b vectors carrying properly cloned HBM nucleotide sequences were picked and incubated in 5 ml of LB broth plus 25 µg/ml kanamycin sulfate overnight. The following day plasmid DNA was isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, CA).

The pET vector can be propagated in any *E. coli* K-12 strain, e.g., HMS174, HB101, 15 JM109, DH5 and the like, for purposes of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts were lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase was induced by addition of isopropyl-β-D-thiogalactoside (IPTG), and the T7 20 RNA polymerase transcribes any target plasmid containing a functional T7 promoter, such as pET-28b, carrying its gene of interest. Strains include, for example, BL21(DE3) (Studier *et al.*, 1990 *Meth. Enzymol.*, 185: 60-89).

To express the recombinant HBM sequence, 50 ng of plasmid DNA are isolated as described above to transform competent BL21(DE3) bacteria as described above (provided 25 by Novagen as part of the pET expression kit). The lacZ gene (β-galactosidase) is expressed in the pET-System as described for the HBM recombinant constructions. Transformed cells were cultured in SOC medium for 1 hour, and the culture was then plated on LB plates containing 25 µg/ml kanamycin sulfate. The following day, the bacterial colonies were pooled and grown in LB medium containing kanamycin sulfate (25 µg/ml) to an optical

density at 600 nM of 0.5 to 1.0 O.D. units, at which point 1 mM IPTG was added to the culture for 3 hours to induce gene expression of the HBM recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria were collected by centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 minutes at 4°C. Pellets were resuspended in 50 ml of cold mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.1 mM EDTA (STE buffer). Cells were then centrifuged at 2000 x g for 20 minutes at 4°C. Wet pellets were weighed and frozen at -80°C until ready for protein purification.

19.1 Chinese Hamster Ovary (CHO) Expression System

Alternatively, HBM and Zmax1 (LRP5) may be expressed in eukaryotic cells. Eukaryotic cells, such as mammalian derived cell lines, are more capable of expressing properly folded proteins containing cystine rich domains such as the EGF and LDLR modules.

19.2 Development of Constructs

HBM and Zmax1 (LRP5) extracellular domain fusions (ECD) to IgG-Fc were prepared. These ECD fusions to the IgG-Fc domain remove the endogenous transmembrane and cytoplasmic portion of the Zmax1/HBM receptor and should produce a secreted fusion protein. The Fc region is separated from the Zmax1/HBM ECD by an enterokinase recognition site so that purified Zmax1 or HBM ECD protein can be obtained without the Fc domain present. The vector used for this construct was pHTop, a derivative of pED (Kaufman *et al.*, 1991 *Nuc. Acids Res.* 19: 4485-90) in which the majority of the adenomajor late promoter was replaced by six repeats of the tet operator (Gossen *et al.*, 1992 *Proc. Natl. Acad. Sci. USA* 89: 5547-51). This vector contains the dihydrofolate reductase (*dhfr*) gene, and when introduced in the cell line CHO/A2 (see description below), functions very efficiently. Clones with high expression can be selected by isolating cells which survive in high methotrexate (MTX) concentrations.

The CHO expression vector pHTOP-Fc was digested with *SalI* and *NotI*. The intervening sequence was purified away from the rest of the vector by electroelution from an

acrylamide gel slice. *SaII* cuts 5' to the intrinsic honey bee mellitin signal sequence in pHTOP-Fc, and *NotI* cuts just 5' to the coding sequence IgG1-Fc. The resulting *SaII-NotI* pHTOP-Fc vector has the signal sequence removed and the *NotI* cloning site is amenable to creating a 5' fusion to IgG-Fc. Full-length Zmax1 (LRP5) in pCMVSPORT6 and full-length HBM in pCMVSPORT6 were digested individually with *XmaI* which cuts within the region of the ORF that encodes the signal sequence) and *BamHI* (that cuts internally in the ORF) to generate a 2286 bp 5' fragment of Zmax1 and HBM. The mutation which distinguishes Zmax1 from HBM lies on this fragment. Separately, the Zmax1 DNA was digested with *BamHI* and *SacI* to isolate an 1800 bp 3' fragment which is common to both the *Zmax1* and the *HBM* genes. Together, these two fragments constitute the coding sequence for the HBM and Zmax1 extracellular domains, less the coding sequence for the first 6 amino acids of the signal sequence and ending 18 amino acids prior to the end of the extracellular domain, which we estimated from Kyte-Doolittle plots to end at the sequence "SPAHSS" (SEQ ID NO:698).

A synthetic duplex was designed to recreate the coding sequence of the Zmax1/HBM signal sequence 5' of the native *XmaI* site, which included the initiator methionine and Kozak sequence. This duplex was designed to contain *SaII* (5') and *XmaI* (3') cohesive ends to adapt ends to adapt the gene fragments described above to the pHTOP-Fc vector. This synthetic duplex was constructed from two partially complementary oligonucleotides as given below:

5'-TCGACCACCATGGAGGCAGCGCCGC-3' (SEQ ID NO:699)

3'-GGTGGTACCTCCGTCGCGGCGGGCC-5' (SEQ ID NO:700)

A second synthetic duplex was designed to recreate the 3' coding sequence from a native *SacI* site to the estimated end of the extracellular domain following the serine in the sequence ". . .SPAHSS", and to also encode a cloning site to allow in-frame fusion to the downstream IgG-Fc. This duplex was designed to contain *SacI* (5') and *NotI* (3') cohesive ends to adapt the gene fragments described above to the pHTOP-Fc vector. This synthetic duplex was constructed from two partially complementary oligonucleotides whose sequences are given below:

5'-CATGTGTGAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAG

CAGTGGC-3' (SEQ ID NO:701)

3'-TCGAGTACACACTTTAGTGGTTCGGCGGGAGTCTGCTGTC GGGCCGGGT
GTCGTCACCGCCGG-5' (SEQ ID NO:702)

5 The fragments, synthetic duplexes, and vector were ligated together in a single reaction. Separate reactions were performed for Zmax (LRP5) 1 and HBM. The ligation mixtures were used to transform electrocompetent *E. coli* DH10B cells, and the resulting colonies were screened by radioactive colony hybridization using the common *SacI*-*Bam*HI 3' fragment as a probe. Colonies containing plasmids with the Zmax1 or HBM fragment inserted were identified, and plasmids were isolated from multiple candidates and their
10 sequences were verified by DNA sequencing. Verified constructs were then used for transfection into CHO cells.

19.2.1 Establishment of CHO stable cell lines

15 The CHO/A2 cell line is derived from CHO DUKX B11 (Urlaub *et al.*, 1980 *Proc. Natl. Acad. Sci. USA* 77: 4216-20) by stably integrating a transcriptional activator (tTA), a fusion protein, between the Tet repressor and the herpes virus VP16 transcriptional domain (Gossen *et al.*) CHO cell lines expressing extracellular HBM-1.Fc and Zmax1.Fc were established by transfecting (using lipofection) pHTopHBM-1.Fc into CHO/A2 cells and pHTopZmax1.Fc into CHO/A2 cells. Clones were selected using by culturing the cells in
20 0.02 μ M methotrexate. Clones were later amplified step-wise to a final concentration of 0.5 μ M methotrexate.

19.2.2 Screening of CHO stable cell lines

25 Multiple clones were screened by a variety of techniques. Clones were screened by Western blot assay using a (mouse) anti-human IgG.Fc horseradish peroxidase (HRP) antibody. The same clones were also metabolically labeled with 35 S-Met/Cys) for a 6 hour pulse, or a 15 minute pulse, followed by a 1 hour, 4 hour, or 24 hour chase in media without radiolabeled Met/Cys. Immunoprecipitations were performed on proteins obtained from conditioned media, as well as from cell extracts. Purification is then performed followed by

sequencing of the proteins using N-terminal sequencing as known in the art.

19.2.3 Fusion Protein Purification

Zmax1-IgG or HBM-IgG fusion protein can be purified from conditioned media or cell extracts of CHO stable cells. The fusion protein is isolated by affinity binding to protein A (for example using protein A coated beads or columns). The IgG-FC domain can then subsequently be cleaved from the Zmax/HBM1 ECD protein by enterokinase digestion.

19.2.4 Potential uses for cell lines and protein

Stable cell lines may be used for generation of purified protein for use in ligand hunting, antibody generation, determination of crystal structure, and competitive binding assays.

A variety of methodologies known in the art can be used to purify the isolated proteins (Coligan *et al.*, *Current Protocols in Protein Science*, John Wiley & Sons (1995)). For example, the frozen cells can be thawed, resuspended in buffer and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International Corp., Newton, MA). The resultant homogenate is centrifuged to yield a clear supernatant (crude extract) and, following filtration, the crude extract is fractioned over columns. Fractions are monitored by absorbance at OD₂₈₀ nm and peak fractions may be analyzed by SDS-PAGE.

The concentrations of purified protein preparations are quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, 1986 *Eur. J. Biochem.* 157:169-180). Protein concentrations are also measured by the method of Bradford, 1976 *Anal. Biochem.* 72: 248-54 and Lowry *et al.*, 1951 *J. Biol. Chem.* 193: 265-275 using bovine serum albumin as a standard.

SDS-polyacrylamide gels of various concentrations were purchased from BioRad (Hercules, CA), and stained with Coomassie blue. Molecular weight markers may include rabbit skeletal muscle myosin (200 kDa), *E. coli* β -galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4

kDa) and bovine aprotinin (6.5 kDa).

Once a sufficient quantity of the desired protein has been obtained, it may be used for various purposes. A typical use is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo* techniques well known in the art. Monoclonal antibodies to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas (Kohler, 1975 *Nature*, 256: 495). In summary, a mouse is inoculated with a few micrograms of HBM protein over a period of two weeks. The mouse is then sacrificed. The cells that produce antibodies are then removed from the mouse's spleen. The spleen cells are then fused with polyethylene glycol with mouse myeloma cells. The successfully fused cells are diluted in a microtiter plate and growth of the culture is continued. The amount of antibody per well is measured by immunoassay methods such as ELISA (Engvall, 1980 *Meth. Enzymol.* 70: 419). Clones producing antibody can be expanded and further propagated to produce HBM antibodies. Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors. See Huse *et al.*, 1989 *Science* 246: 1275-81. For additional information on antibody production see Davis *et al.*, Basic Methods in Molecular Biology, Elsevier, NY, Section 21-2 (1989).

Additional uses for purified or isolated protein includes use in X-ray crystallography, binding assays, and so forth as described in greater detail *infra*.

19.3 Zmax1, LRP6, and Variant Antibodies

Polyclonal antibodies were developed to both human Zmax1 (LRP5) (SEQ ID NO:3) and LRP6 (GenBank Accession No. AF074264). Antibodies can similarly be prepared against HBM and HBM like proteins and polypeptides. Peptides from the Zmax1 amino acid sequence were selected as immunogens based on five goals. 1) Maximize differences between Zmax1 and LRP6 amino acid sequences (71% amino acid identity). See Fig. 27. For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference

sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s), relative to the reference sequence, based on the designated program parameters. 2) Minimize potential cross reactivity with other known genes by performing sequence alignment and similarity searches. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith *et al.*, 1981 *Adv. Appl. Math.* 2: 482, by the homology alignment algorithm of Needleman *et al.*, 1970 *J. Mol. Biol.* 48: 443, by the search for similarity method of Pearson *et al.*, 1988 *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms and others in programs contained in the Wisconsin genetics software package, Genetics Computer Group, 585 Science Dr., Madison, WI, or by visual inspection (see generally Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997). Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, 1990 *J. Mol. Biol.* 215: 403-10. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. 3) Obtain peptides with the highest antigenicity index as possible as determined by PeptideStructure protein analysis using software programs contained in the Wisconsin Genetics software package, Genetics Computer Group, 575 Science Dr., Madison, WI. 4) Locating peptides relative to highly homologous domains (e.g., EGF-like domains and LDL receptor repeats) within the gene family and the location relative to the extracellular and cytoplasmic regions of the gene. 5) And, for human Zmax1 specific antibodies, the human amino acid sequence (SEQ ID NO: 3) was compared to the mouse Zmax1 sequence (GenBank Accession No. AF064984) and peptides were selected based on the above criteria in addition to minimizing the sequence similarity between the two species (See Figure 26).

Using the same criteria above, LRP6 specific peptides were selected for polyclonal antibody production. Table 11 lists the amino acid sequences that were chosen, the amino acid differences within the peptide between the human and mouse sequences. All the peptide sequences (ranging from 12-18 amino acids) were provided to Sigma/Genosys (St. Louis,

MO) for peptide synthesis and subsequent polyclonal antibody production in New Zealand White Rabbits. The IgG fraction from the serum of each immunized rabbit was isolated using Protein G Sepharose (Amersham). Polyclonal antibody generation using these peptides may be done in other species as well, for example, chickens. This is often advantageous when there is a high degree of similarity between the human (reference) and murine/rodent sequence.

TABLE 11

Amino Acids	Amino Acids	SEQ ID NO	H/M* Differences	Comments
171-187	VETPRIERAGMDGSTRK	703	5	Contains HBM polymorphism
264-278	NKRTGGKRKEILSAL	704	3	Extracellular
290-301	ERQPFHTRCEE	705	2	Adjacent to EGF-I, extracellular
532-546	VDGTKRRTLLEDKLP	706	5	Extracellular
901-915	DGLNDCMHNNQCGQ	707	2	In EGF-III, extracellular
1010-1021	PFVLTSLSQGQN	708	6	Extracellular, human specific
1415-1429	YAGANGPFPHEYVSG	709	3	Cytoplasmic
1452-1464	ACGKSMMSSVSLM	710	5	Cytoplasmic, human specific
1556-1573	RWKASKYYLDLNSDSPY	711	1	Cytoplasmic
888-902	SGWNECASSNGHCSH	712		LRP6 specific
1308-1321	NGDANCQDKSDEKN	713		LRP6 specific

* H/M-differences between human and mouse sequences

Antibodies towards variants of LRP6, HBM and LRP5 are also contemplated. Based on the analysis of the structural model of the LRP5 beta propeller 1 (discussed in more detail in the Examples), interior regions of the propeller were analysed. Since the site-directed mutagenesis experiments had confirmed that modulation of propeller 1, particularly in the interior regions of beta propeller 1, could result in an HBM effect, a strategy was employed to generate antibodies with epitopes specific to these regions. Such antibodies to the wild type

LRP5 receptor could serve, for examples, as an HBM mimetic, by altering ligand/receptor interactions, protein-protein interactions, or by modulating Wnt or Dkk activity. Such antibodies could be used as therapies to treat, for example osteoporosis.

One set of preferred antibodies includes:

5	LRP5 Amino Acids	Sequence
	208-223	KLYWADAKLSFIHRAN
	277-291	ALYSPMDIQVLSQER
	61-73	GLEDAAAVDFQFSKGA
	234-247	EGSLTHPFALTLSG
10	249-264	TLYWTDWQTRSIHACN
	144-156	VLFWQDLDQPRAI
	194-210	IYWPNGLTIDLEEKLY
	34-47	LLLFANRRDVRLVD
	75-89	GAVYWTDVSEEAIKQ
15	121-135	KLYWTDSETNRIEVA

Similar peptides can be used to prepare antibodies based on the LRP6 or HBM structures and the proposed mutations in their propellers. Other antibodies could easily be prepared based on the structural model discussed in the Examples, as would be readily appreciated by the skilled artisan.

20 Another series of antibodies could be prepared which target the domains on HBM, LRP5, LRP6 and their variants which interact Dkk. Such antibodies could again serve as HBM mimetics, for example by displacing Dkk binding and thereby could be used as an osteoporosis therapeutic. Preferred peptides for preparing antibodies include but are not limited to:

25	LRP5 Amino Acids	Sequence
	969-993	LILPLHGLRNVKAIDYDPLDKFIYW
	989-1013	KFIYWVDGRQNIKRAKDDGTQPFVL

LRP5 Amino Acids Sequence

1009-1033 QPFVLTSLSQGQNPDRQPHDLSIDI
 1029-1053 LSIDIYSRTLFWTCEATNTINVHRL
 10491073 NVHRLSGEAMGVVLRGDRDKPRAIV
 1253-1266 CGEPPTCSPDQFAC
 1278-1295 WRCDGFPECDDQSDEEGC
 1316-1332 RCDGEADCQDRSDEADC
 1370-1383 CEITKPPSDDSPA

Additional peptides would be readily apparent to the artisan of ordinary skill.

19.3.1 Single chain Fv molecules developed by phage display

Peptides were chosen from the Zmax1 (LRP5) sequence (SEQ ID NO: 3) to screen for single chain Fv (scFv) molecules by phage display. Similar peptides can be chosen for LRP6, HBM and HBM-like proteins to screen for scFv molecules. As discussed below, all mention of LRP5 is also meant to include these other proteins.

A total of 17 peptides from the Zmax1 sequence were selected for synthesis and subsequent phage display screen for scFv molecules. All peptide synthesis and phage display work was performed at Cambridge Antibody Technology (CaT) in Cambridge, UK.

Peptides were selected based on criteria as described above.

TABLE 12

Protein Domain	Zmax1 Residues	LRP6 Residues	% Identity
Spacer 1 (+G171V)	161-181	148-168	76%
Spacer 1 (-G171V)	161-181	148-168	76%
EGF 1	301-321	288-308	76%
Spacer 2	401-421	388-408	52%
EGF 2	611-631	598-618	62%
Spacer 3	781-801	768-788	62%

Protein Domain	Zmax1 Residues	LRP6 Residues	% Identity
EGF 3	921-941	908-928	10%
Spacer 4	1000-1021	988-1008	26%
EGF 4	1229-1249	1219-1239	76%
LDLR 1	1261-1282	1252-1272	81%
LDLR 2	1300-1320	1290-1310	57%
LDLR 3	1338-1358	1328-1348	48%
Cytosolic 1	1418-1438	1405-1425	14%
Cytosolic 1	1516-1536	1503-1525	52%
Cytosolic 1	1535-1555	1524-1544	81%
Cytosolic 1	1595-1615	1592-1613	82%
Spacer 2 (cross reactive)	421-441	408-428	100%

Note that a number of these regions (e.g., 401-421, 421-441, 781-801, and 1229-1249) share 100% identity with mouse LRP5 (see Fig. 26). Therefore, these may be used against both mouse and human forms of the protein. The peptide 421-441 was included to facilitate the generation of an antibody that would recognize both Zmax1 (LRP5) and LRP6 (see Figure 27). Two peptides were synthesized spanning the HBM mutation site (Zmax1 residues 161-181), one with the Zmax1 sequence and the other containing the HBM sequence.

Once scFv molecules were isolated, they were used as reagents in immunochemistry to detect Zmax1 (LRP5) protein expression in a variety of human normal and diseased tissues. The details of the scFV antibody immunohistochemical analysis of three phage clones against peptide 1000-1021 (i.e., IKRAKDDGTQPFVLTSLSQGQN; SEQ ID NO: 714) of the extracellular domain of Zmax1 showed positive staining with cardiac muscle, kidney, lung and liver. Expression was also detected in prostate carcinoma. These results are consistent with mRNA tissue distribution profiles as well as with the published reports of LRP5 mRNA localization (Kim *et al.*, 1998 *J. Biochem.* 124: 1072-6). The resulting phage clones arise from pools and will be sequenced to identify potential variants in the Fv

region of the molecules. Once identified, the suitable scFVs can then be subcloned into variable heavy chain and variable light chain DNA constructs for cotransfection into COS cells for final assembly of an intact and functional immunoglobulin gamma (IgG) molecule. The IgG that is expressed by the cells can then be further characterized for specificity and reactivity as would be known in the art.

19.3.2 Monoclonal antibody development

Monoclonal antibodies can be developed to Zmax1 (LRP5), LRP6, HBM and HBM-like proteins/polypeptides by complete cell and adenovirus immunization of, for example, Balb/c mice (antibodies to all forms are contemplated even when only a few examples are discussed in detail). Dendritic cells can be isolated from spleens of Balb/c mice, for example, and the cells expanded *in vitro* in the presence of growth factors IL-4 and GM-CSF. The dendritic cells can then be infected with HBM or Zmax1 adenovirus particles. The cells are then cultured for 24 hours prior to intravenous injection into Balb/c mice. Dendritic cells (1×10^6 cells/mouse) are injected 2-3 times every 3-4 weeks over a three month period.

Alternatively, purified HBM and Zmax DNAs in, for example, the pcDNA3.1 expression vector, can be coated on colloidal gold particles. These particles can then be injected subcutaneously into the desired mouse using gene gun technology. Approximately, $5 \mu\text{g}$ cDNA/mouse can be injected. Injections are performed 4-6 times every 2 weeks over approximately a 3 month period.

Another option is that cells (any species of animal, but preferably Balb/c mouse strain or the same species as the mouse strain being used which is related to limit antigen response to non-specific protein) over expressing HBM and Zmax1 (LRP5) and their respective adenovirus will be injected into the mice every 2-3 weeks for a period of about 1.5 to 3 months, as necessary. The bleeds from the mice can be tested for reactivity with the native and denatured protein by ELISA (using purified protein or protein-fusions), cell based ELISA, immunohistochemical staining and Western blotting. Serum samples from

the animals can be screened by FACS (fluorescent activated cell sorting) using cells infected with Zmax1 or HBM adenovirus. The spleen cells (antibody producing cells) from the mice with the strongest reactivity can then be fused with a myeloma to generate the hybridoma cells. The conditioned media from the hybridoma is then screened for the positive cell colonies for subsequent cloning. These cloned cells can then be injected into the intraperitoneal space in mice for ascites production.

19.3.3 Polyclonal Antibody Applications

Polyclonal antibodies directed against Zmax1 (LRP5) and LRP6 were developed to determine the function of these proteins, analyze the expressed pattern and levels in various tissues, cells or any biological sample. Similar antibodies could be prepared which distinguish the wild type forms from the HBM and HBM-like variants.

Uses for polyclonal antibodies against Zmax1, HBM, HBM-like polypeptides, and LRP6 include but are not limited to: analysis of bone cross-sectional mounts, tissue distribution, evaluation of expression of the protein from bone biopsy samples of affected/non-affected family members (e.g., bone cell digests, explants of bone marrow stromal cell cultures), evaluation of protein expression levels in transiently or stably transfected cells, evaluating protein concentration in tissues, serum or body fluid, purification of full length or fragments of these proteins for ligand hunting and functional assay development, identification of ligands or proteins which interact with these proteins, and elucidations of the signaling pathways of LRP6, Zmax1 (LRP5) and HBM, and related variants.

For example, Zmax1 cloned in pcDNA3.1 (Invitrogen, Carlsbad, CA). This was used to generate ³⁵S-labeled *in vitro* translated (Promega, Madison, WI) Zmax1. Antibody (10 µg/ml) 3109 and 3110, which are directed against peptide immunogen, RWKASKYYLDLNSDSDPY (SEQ ID NO:711), was combined with 20 µl of the *in vitro* translated product in the presence of either 10 µg/ml specific peptide (*i.e.*, RWKASKYYLDLNSDSDPY; SEQ ID NO: 711) or non-specific peptide (*i.e.*,

SGWNECASSNGHCSH; SEQ ID NO: 712) or no peptide and incubated for 1.5 hr at 4°C. Protein A Sepharose was then added to the samples (previously blocked for about 1.5 hr with reticulocyte lysate), and the samples were shaken for 1 hour at 4°C. The protein A Sepharose was washed 3 times with 0.5 ml of PBS. The bound protein was subsequently separated on a 4-12% gradient NuPAGE gel (Invitrogen) according to manufacturer's instructions. The gel was dried at 80°C for 30 min and then exposed to Kodak X-OMAT-AR film for 24 to 48 hr. The specific peptide was observed to significantly compete for the ³⁵S-labeled Zmax1 immunoprecipitated protein with either antibody. The competition was not observed with a non-specific peptide.

These antibodies can also be used for immunohistochemistry. For example, HBM transgenic and wild-type mice were sacrificed using CO₂ narcosis. Mouse calvariae were removed intact, and the soft tissues gently dissected. The bones were fixed in 10% phosphate buffered formalin for 24 hours for further processing and analysis. After fixation, calvariae were decalcified in TBD-2 decalcifying agent (Shandon, Pittsburgh, PA) for about 7-8 hours and then dehydrated in graded alcohol. Calvariae were then bisected perpendicular to the sagittal suture through the central portion of the parietal bones parallel to the lambdoidal and coronal sutures and embedded in paraffin. Four to six 5 μm thick representative sections were cut.

For example, the rabbit polyclonal antibody, Zmax1/HBM (i.e., antibody 3109 and 3110) recognize Zmax1 (LRP5) in both HBM transgenic and wild-type mouse calvariae. An anti-Zmax1 or anti-HBM antibody can be used to detect Zmax1 or HBM protein in order to evaluate its abundance and pattern of protein expression. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, P-galactosidase, or acetylcholinesterase; example of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine

fluorescein, dansyl chloride or phycoerythrin; and example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin and acquorien; and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S , and ^3H . Alternatively, a secondary antibody can be employed that detects the presence of the primary Zmax1 polyclonal antibody. An example would be an antibody that recognized all rabbit immunoglobulins. This secondary antibody could be coupled in an identical manner as described above to facilitate detection. Controls comprised samples with the avidin peroxidase, but without antibody. Intensive positive staining of stroma cells and mesenchymal cells was observed in the suture area. Pre-osteoblasts and osteoblasts were observed to stain within the periosteum and some osteocytes with antibody 3109 and 3110 in the calvariae of the HBM compared to wild-type mice. High magnification of tissue calvaria sections of the HBM transgenic mice showed a pronounced cell membrane staining of the osteocytes and the cells within the suture area.

Similar antibodies could be prepared to HBM-like proteins and polypeptides, as would be readily appreciated by the artisan of ordinary skill.

20. Methods of Use: Gene Therapy

In recent years, significant technological advances have been made in the area of gene therapy for both genetic and acquired diseases (Kay *et al.*, 1997 *Proc. Natl. Acad. Sci. USA*, 94: 12744-6). Gene therapy can be characterized as the deliberate transfer of DNA for therapeutic purposes. Improvement in gene transfer methods has allowed for development of gene therapy protocols for the treatment of diverse types of diseases. Gene therapy has also taken advantage of recent advances in the identification of new therapeutic genes, improvement in both viral and nonviral gene delivery systems, better understanding of gene regulation, and improvement in cell isolation and transplantation.

The preceding experiments identify the *HBM* gene as a dominant mutation conferring elevated bone mass. Additional HBM-like genes are identifiable based on the model and data presented herein regarding the propellers and structure of the HBM protein. The fact that this mutation is dominant indicates that expression of the HBM protein causes elevated bone

mass. Older individuals carrying the *HBM* gene, and, therefore expressing the HBM protein, do not suffer from osteoporosis. These individuals are equivalent to individuals being treated with the HBM protein. These observations are a strong experimental indication that therapeutic treatment with the HBM protein prevents osteoporosis. The bone mass elevating activity of the *HBM* gene is termed "HBM function" or "HBM like phenotype" in the instance of other HBM variants.

Therefore, according to the present invention, a method is also provided of supplying HBM function to mesenchymal stem cells (Onyia *et al.*, 1998 *J. Bone Miner. Res.* 13: 20-30; Ko *et al.*, 1996 *Cancer Res.* 56: 4614-9). Supplying such a function provides protection against osteoporosis. The *HBM* gene or a part of the gene or other *HBM*-like gene may be introduced into the cell in a vector such that the gene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location.

Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for introducing DNA into cells such as electroporation, calcium phosphate co-precipitation, and viral transduction are known in the art, and the choice of method is within the competence of one skilled in the art (Robbins, ed., *Gene Therapy Protocols*, Human Press, NJ (1997)). Cells transformed with the *HBM* gene (or HBM-like gene) can be used as model systems to study osteoporosis and drug treatments that promote bone growth.

As generally discussed above, the *HBM* or HBM-like gene or biologically active fragment thereof, where applicable, may be used in gene therapy methods in order to increase the amount of the expression products of such genes in mesenchymal stem cells (in all instances where discussing HBM gene and its cognate product, the HBM-like gene and cognate protein is also contemplated). It may be useful also to increase the level of expression of a given HBM protein, or a fragment thereof, even in those cells in which the wild type gene is expressed normally. Gene therapy would be carried out according to generally accepted methods as described by, for example, Friedman, *Therapy for Genetic Diseases*, Friedman, Ed., Oxford University Press, pages 105-121 (1991).

A virus or plasmid vector containing a copy of the *HBM* gene linked to expression

control elements and capable of replicating inside mesenchymal stem cells, is prepared.

Suitable vectors are known and described, for example, in U.S. Patent No. 5,252,479 and WO 93/07282, the disclosures of which are incorporated by reference herein in their entirety. The vector is then injected into the patient, either locally into the bone marrow or systemically (in order to reach any mesenchymal stem cells located at other sites, i.e., in the blood). If the transfected gene is not permanently incorporated into the genome of each of the targeted cells, the treatment may have to be repeated periodically.

Gene transfer systems known in the art may be useful in the practice of the gene therapy methods of the present invention. These include viral and non-viral transfer methods.

A number of viruses have been used as gene transfer vectors, including polyoma, i.e., SV40 (Madzak *et al.*, 1992 *J. Gen. Virol.* 73: 1533-6), adenovirus (Berkner, 1992 *Curr. Top. Microbiol. Immunol.* 158: 39-61; Berkner *et al.*, 1988 *BioTechniques*, 6: 616-629; Gorziglia *et al.*, 1992 *J. Virol.* 66: 4407-12; Quantin *et al.*, 1992 *Proc. Natl. Acad. Sci. USA* 89: 2581-2584; Rosenfeld *et al.*, 1992 *Cell* 68: 143-155; Wilkinson *et al.*, 1992 *Nucl. Acids Res.* 20: 2233-39; Stratford-Perricaudet *et al.*, 1990 *Hum. Gene Ther.* 1: 241-256), vaccinia virus (Mackett *et al.*, 1992 *Biotechnology* 24: 495-499), adeno-associated virus (Muzyczka, 1992 *Curr. Top. Microbiol. Immunol.* 158: 91-123; Ohi *et al.*, 1990 *Gene* 89: 279-282), herpes viruses including HSV and EBV (Margolskee, 1992 *Curr. Top. Microbiol. Immunol.* 158: 67-90; Johnson *et al.*, 1992 *J. Virol.*, 66: 2952-65; Fink *et al.*, 1992 *Hum. Gene Ther.* 3: 11-9; Breakfield *et al.*, 1987 *Mol. Neurobiol.* 1: 337-371; Fresse *et al.*, 1990 *Biochem. Pharmacol.*, 40: 2189-2199), and retroviruses of avian (Brandyopadhyay *et al.*, 1984 *Mol. Cell Biol.*, 4: 749-754; Petropoulos *et al.*, 1992 *J. Virol.* 66: 3391-3397), murine (Miller, 1992 *Curr. Top. Microbiol. Immunol.* 158: 1-24; Miller *et al.*, 1985 *Mol. Cell Biol.* 5: 431-437; Sorge *et al.*, 1984 *Mol. Cell Biol.* 4: 1730-7; Mann *et al.*, 1985 *J. Virol.* 54: 401-7), and human origin (Page *et al.*, 1990 *J. Virol.* 64: 5370-6; Buchsachalcher *et al.*, 1992 *J. Virol.* 66: 2731-9). Most human gene therapy protocols have been based on disabled murine retroviruses.

Non-viral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham *et al.*, 1973 *Virology* 52: 456-67; Pellicer *et al.*, 1980 *Science* 209: 1414-22), mechanical techniques, for example microinjection (Anderson

et al., 1980 *Proc. Natl. Acad. Sci. USA*, 77: 5399-5403; Gordon *et al.*, 1980 *Proc. Natl. Acad. Sci. USA*, 77: 7380-4; Brinster *et al.*, 1981 *Cell* 27: 223-231; Constantini *et al.*, 1981 *Nature* 294: 92-94), membrane fusion-mediated transfer via liposomes (Felgner *et al.*, 1987 *Proc. Natl. Acad. Sci. USA* 84: 7413-7; Wang *et al.*, 1989 *Biochemistry* 28: 9508-14; Kaneda *et al.*, 1989 *J. Biol. Chem.* 264: 12126-9; Stewart *et al.*, 1992 *Hum. Gene Ther.* 3: 267-275; Nabel *et al.*, 1990 *Science* 249: 1285-8; Lim *et al.*, 1992 *Circulation* 83: 2007-11), and direct DNA uptake and receptor-mediated DNA transfer (Wolff *et al.*, 1990 *Science* 247: 1465-8; Wu *et al.*, 1991 *BioTechniques* 11: 474-85; Zenke *et al.*, 1990 *Proc. Natl. Acad. Sci. USA* 87: 3655-9; Wu *et al.*, 1989 *J. Biol. Chem.* 264: 16985-7; Wolff *et al.*, 1991 *BioTechniques* 11: 474-45; Wagner *et al.*, 1990; Wagner *et al.*, 1991 *Proc. Natl. Acad. Sci. USA*, 88: 4255-9; Cotten *et al.*, 1990 *Proc. Natl. Acad. Sci. USA*, 87: 4033-7; Curiel *et al.*, 1991 *Proc. Natl. Acad. Sci. USA*, 88: 8850-4; Curiel *et al.*, 1991 *Hum. Gene Ther.* 3: 147-54). Viral-mediated gene transfer can be combined with direct *in vivo* vectors to the mesenchymal stem cells and not into the surrounding cells (Romano *et al.*, 1998 *In Vivo*, 12: 59-67; Gonez *et al.*, 1998 *Hum. Mol. Genetics* 7:1913-9). Alternatively, the retroviral vector producer cell line can be injected into the bone marrow (Culver *et al.*, 1992 *Science* 256: 1550-2). Injection of producer cells would then provide a continuous source of vector particles. This technique has been approved for use in humans with inoperable brain tumors.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

Liposome/DNA complexes have been shown to be capable of mediating direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is non-specific, localized *in vivo* uptake and expression have been reported in tumor deposits, for example, following direct *in situ* administration (Nabel, *Hum. Gene Ther.*, 3:399-410 (1992)).

21. Methods of Use: Transformed Hosts and Transgenic Animals as Research Tools and for the Development of Pharmaceuticals

Cells and animals that carry the *HBM*, *HBM*-like, *Zmax1* (*LRP5*), or *LRP6* genes, used as model systems, are valuable research tools to study and test for substances that have potential as therapeutic agents (Onyia *et al.*, 1998 *J. Bone Miner. Res.*, 13: 20-30; Broder *et al.*, 1997 *Bone*, 21: 225-35). Discussion of one of these is meant to include the others, e.g., discussion of *HBM* is meant to include *HBM*-like variants, and so forth.

Cells for this purpose are typically cultured mesenchymal stem cells. These may be isolated from individuals with somatic or germline *HBM* genes. Alternatively, the cell line can be engineered to carry the *HBM* or *HBM*-like gene, as described above. After a test substance is applied to the cells, the transformed phenotype of the cell is determined. Any trait of transformed cells can be assessed, including, for example, formation of bone matrix in culture (Broder *et al.*, 1997 *Bone*, 21: 225-235), mechanical properties (Kizer *et al.*, 1997 *Proc. Natl. Acad. Sci. USA* 94: 1013-8), expression of marker genes and response to application of putative therapeutic agents.

Transgenic modified animals and cell lines may be used to test therapeutic agents. Transgenic modifications include, for example, insertion of the *Zmax1* (*LRP5*) gene, *LRP6*, *HBM* gene or *HBM*-like gene and disrupted homologous genes. Alternatively, the inserted *Zmax1*, *LRP6*, *HBM*, and/or *HBM*-like gene of the animals may be disrupted by insertion or deletion mutation of other genetic alterations using conventional techniques, such as those described by, for example, Capecchi, 1989 *Science* 244: 1288; Valancuis *et al.*, 1991 *Mol. Cell Biol.* 11: 1402; Hasty *et al.*, 1991 *Nature* 350: 243; Shinkai *et al.*, *Cell*, 68:855 (1992); Mombaerts *et al.*, *Cell*, 68:869 (1992); Philpott *et al.*, 1992 *Science* 256: 1448; Snouwaert *et al.*, 1992 *Science* 257: 1083; Donehower *et al.*, 1992 *Nature* 356: 215. After test substances have been administered to the animals, the growth of bone must be assessed. If the test substance modulates (e.g., enhances) the growth of bone, then the test substance is a candidate therapeutic agent. These animal models provide an extremely important vehicle for potential therapeutic products.

The present invention also provides animals and cell lines wherein the expression of

endogenous genes are activated, and may be further amplified, which does not require *in vitro* manipulation and transfection of exogenous DNA encoding Zmax1 (LRP5) or HBM proteins. These methods are described for example in PCT Application WO 94/12650 and U.S. Patent No. 5,968,502, both of which are herein incorporated in their entirety by reference. In addition, the present invention includes methods wherein endogenous activation or over-expression is achieved by *in situ* homologous recombination, non-homologous recombination, or illegitimate recombination methods. These methods are described for example in PCT Applications WO 99/15659 and WO 00/49162, both of which are incorporated herein in their entirety.

21.1 Creating transgenic and gene-targeted animals

The present invention provides genetically modified animals that recapitulate the human HBM or a HBM-like phenotype. The approaches taken involve the creation of both transgenic and gene-targeted animals that have the human G to T nucleotide substitution incorporated into the genome, express human Zmax1 (LRP5) or express a variant with a bone mass altering phenotype. These approaches can be used with any gene, such as HBM, HBM-like, LRP5 and LRP6 systems.

21.1.1 Transgenic mice over-expressing the HBM polymorphism

Plasmid constructs were prepared that utilized either the CMVbActin or type I collagen promoters to drive expression of the human HBM cDNA. The most commonly-used promoters for mammalian expression are from cytomegalovirus (CMV), Rous sarcoma virus (RSV), Simian virus 40 (SV40), and EF-1a (human elongation factor 1a-subunit). CMV is derived from the human cytomegalovirus immediate-early viral promoter. CMV is a stronger promoter in most cell lines than either RSV or SV40. The RSV promoter is derived from an avian virus and tends to be a strong promoter in avian cell lines. The SV40 promoter expresses well in cell lines that carry the large T antigen, such as COS-1. In these cell lines, the plasmid is replicated to higher copy numbers. EF-1a is beginning to be more widely used for recombinant protein expression. EF-1a is the promoter from the human

elongation factor 1a-subunit, a gene that is highly expressed and conserved in mammalian cell lines.

The chimeric CMVbActin promoter is a strong promoter that has been shown to produce ubiquitous gene expression in transgenic mice including bone. This promoter was chosen to drive expression of HBM in a manner consistent with the reported widespread expression of the endogenous mouse *Zmax1* (*LRP5*) gene. Although the HBM phenotype is observed in bone, the *HBM* gene may have direct or indirect effects in other tissues. Therefore, other strong ubiquitous promoters may be utilized as would be known to those skilled in the art.

Type I collagen promoters provide tissue-restricted gene expression wherein expression is primarily limited to bone. Other bone-specific promoters are available that could result in expression of HBM in bone. For example, promoters associated with proliferation of osteoblasts include histone, type I collagen, TGF β 1, MSX2, cfos/cJun and Cbfa1 may be used. Promoters associated with bone matrix maturation including alkaline phosphatase, MGP, Cbfa1, Fra/Jun and Dlx5 also can be used. Promoters associated with bone mineralization such as osteocalcin, osteopontin, bone sialoprotein and collagenase also can be used. The promoter chosen would be determined by, for example, the tissue expression, the degree of regulatable control and the like as would be known to the skilled artisan. For example, the type I collagen promoters were chosen to insure that HBM would be expressed in bone in a temporal, spatial and bone cell-specific pattern resembling the endogenous pattern of *Zmax1* (*LRP5*) expression in bone.

21.1.2 Transgenic mice over-expressing the wild-type (*LRP5*) *Zmax1* gene

Plasmid constructs were prepared using the CMVbActin and type I collagen promoters driving expression of *Zmax1* (*LRP5*). These animals can serve as a control animal model for the HBM, HBM-like and LRP6 transgenic mice (all of which are contemplated when discussing LRP5/*Zmax1*). Additional controls include non-transgenic littermates and wild-type animals of an identical genetic background. Methods for preparing these animals would be similar to what is discussed for mice which over express the HBM polymorphism.

21.1.3 *LRP6* gene targeted knock out mice

LRP6 knock-out mice were generated using Omnibank embryonic stem (ES) cells carrying a gene trap vector which inserted into the first intron of the *LRP6* gene. The insert location was determined to be the *LRP6* gene by an Omnibank Sequence Tag (OST) generated by reverse transcription PCR (RT-PCR) of a fusion transcript comprised of 5' gene trap vector sequence spliced to the host gene transcript 3' of the insertion site. The gene trap vector functionally knocks out the mouse *LRP6* gene by forced splicing of *LRP6* exon 1 to the IRES-LacZ-PolyA element of the gene trap, preventing transcription of *LRP*.

Chimeric mice were generated with ES cells, identified as OST38808, by injection into C57BL/6 albino host blastocysts which were then transferred to pseudopregnant females and allowed to develop through birth. Germ line chimeras were backcrossed to 129SvEvBrd strain mice to maintain the knockout allele of *LRP6* on an inbred 1298SvEvBrd genetic background. Germ line transmission of the *LRP6*-KO allele was identified by PCR amplification of a gene trap specific sequence. Heterozygous *LRP6*-KO mating pairs were used for continued breeding. The genotype of wt and *LRP6*-KO progeny is determined by tail DNA PCR.

Measurements of bone density at 9 weeks of age in female heterozygous knock-out mice has shown significant ($p < 0.05$) decreases in bone volume, trabecular number, and trabecular thickness as measured by uCT. These results are consistent with the hypothesis that *LRP6* is also involved in modulating bone density and is a target for development of therapies and drugs. Accordingly, *LRP6* transgenic animals and transgenic animals expressing bone modulating variants of *LRP6* are contemplated within the scope of the invention.

21.1.4 Gene-targeted mice expressing the HBM polymorphism

A gene-targeting construct was prepared that could be used to create animals containing a HBM or HBM-like knock-in (KI) allele and a *Zmax1* (*LRP5*) knock-out (KO) allele. The gene-targeting construct contained the HBM polymorphism in exon 3 and included a neomycin selection cassette that was linked to a transcriptional stop sequence and

was flanked with lox P sites. The HBM polymorphism in mouse *Zmax1* results in a glycine to valine change in the amino acid sequence at position 170 of the mouse LRP5 homolog (Genbank Accession No. AF064984). Homologous recombination is used to stably introduce the construct into the mouse genome. If the transcriptional stop sequence functioned to completely block transcription of the modified *Zmax1* allele, then a functional *Zmax1* knock-out allele would be generated. This would facilitate production of a homozygous knock-out animal for the *Zmax1* gene.

To create the knock-in allele, Cre recombinase could be used to excise the neomycin selection cassette leaving behind the modified exon 3 and one copy of the loxP site. Cre could be introduced into single-cell fertilized embryos to facilitate ubiquitous expression of HBM or by crossing animals with transgenic mice to obtain bone-specific HBM expression. Homozygous animals could be made for the HBM knock-in allele. Alternatively, animals could be created by nuclear transfer techniques, wherein nuclei from homozygous animals is transferred into a prepared oocyte (e.g., enucleated) as is known in the art. See, e.g., Campbell *et al.*, *Nature* 380: 64-68 (1996). Additional methods of creating knock out mice include engineering a homologous recombination vector wherein the ATG start codon is deleted or mutated, engineering a frame-shift mutation into the vector, engineering deletions of critical portions of the promoter region, and/ or engineering a vector to delete critical regions of the gene.

The methods used for the LRP5, LRP6 and HBM mice can similarly be used for other variants of these genes.

21.2 Materials and Methods

21.2.1 Construction of the *Zmax1* (LRP5) plasmid *Zmax1GI_3AS*

The full-length *Zmax1* cDNA construct has been engineered into the *XbaI*-*NotI* sites of the pCMVSPORT6.0 vector from Life Technologies (part of the Gateway cloning system) to create *Zmax1GI_3AS*. The insert (5,278 bp) can be released from the vector by digestion with either *HindIII* or *XbaI* on the 5' side together with either *EcoRV* or *EcoRI* on the 3' end. A similar construct can be prepared for LRP6 and HBM-like nucleic acids as would be

apparent to one of ordinary skill and is contemplated as well in the discussion of this section.

The *Zmax1* (LRP5) construct was generated from four independent partial clones. These clones were isolated from a *Zmax1* specific primed cDNA library. A partial *Zmax1* cDNA clone existed in the internal survey sequencing clone set as L236B_P0049E08 isolated from an oligo-dT primed HeLa cell cDNA library. This clone was truncated at the 5-primed end. In order to isolate more 5-prime containing fragments necessary to generate a full length cDNA, a *Zmax1* gene-specific cDNA library was generated from Clontech human liver poly-A mRNA (catalog #6510-1, lot# 9060032) and Life Technologies SuperScript® Plasmid System for cDNA Synthesis and Plasmid Cloning kit (catalog no. 18248-013). This library was designated as L401. The manufacturer's protocol was carried out with the following modifications. 1) In both first and second strand synthesis reactions, DEPC-treated water was substituted for α -³²P-dCTP. 2) Reverse transcription was primed using oligonucleotides that were selected to be specific for the *Zmax1* gene at approximately 1 kb intervals. These sequences were checked using the program BLAST against the public databases to ensure *Zmax1* specificity. 3) Two separate reverse transcription reactions were performed. The first reaction, (A), was primed with oligonucleotides which annealed to the more 3' regions of *Zmax1* (LRP5) as follows:

47114: 5'-CGTACGTAAAGCGGCCGCTTGGCAATACAGATGTGGGA-3' (SEQ ID NO: 715)

47116: 5'-CGTACGTAAAGCGGCCGCGAGTAGCTCCTCTCGGTGGC-3' (SEQ ID NO: 716)

47118: 5'-CGTACGTAAAGCGGCCGCGCTCATCATGGACTTCCG-3' (SEQ ID NO: 717)

47120: 5'-CGTACGTAAAGCGGCCGCGCACTGCTGTTTGATGAGG-3' (SEQ ID NO: 718)

The second reaction, (B), used the previously mentioned four oligonucleotides, as well as

47108: 5'-CGTACGTAAAGCGGCCGCGAGTGTGGAAGAAAGGCTGC-3' (SEQ ID NO: 719)

47110: 5'-CGTACGTAAAGCGGCCGCGAGTAGAGCTTCCCCTCCTGC-3' (SEQ ID NO: 720)

47112: 5'-CGTACGTAAAGCGGCCGCGTCCATCACGAAGTCCAGGT-3' (SEQ ID NO: 721)

All oligonucleotides contained a *NotI* linker sequence and were used at a concentration of 0.02 ug/ul. 4) The *SaII*-adapted cDNA from both reverse transcription reactions was size-fractionated by electrophoresis on 1% agarose, 0.1 ug/ml ethidium bromide, 1x TAE gels.

The ethidium bromide-stained cDNA between 0.6 and 8.0 kb was excised from the gel. The cDNA was purified from the agarose gel by electroelution (ISCO Little Blue Tank Electroelutor[®]) using the manufacturer's protocol. The purified cDNA from reactions A and B were then pooled together. 5) The size-fractionated *SaII*-adapted cDNA was ligated to

5 *NotI*-*SaII* digested pBluescript[®] (Stratagene, La Jolla, CA).

Ligated library cDNA (3 ml) was used to transform electrocompetent *E. coli* cells (ElectroMAX[®] DH10B cells and protocol, Life Technologies catalog no. 18290-015, BioRad *E. coli* pulser, voltage 1.8 KV, 3-5 msec pulse). The transformed cells were plated on LB-ampicillin (100 µg/ml) agar plates and incubated overnight at 37°C. Approximately 10⁶

10 colony forming units (cfu) were plated at a density of 50,000 cfu/150 mm plate. Cells were washed off the plates with LB media, and collected by centrifugation. Plasmid DNA was purified from the cells using the QIAGEN Plasmid Giga Kit and protocol (catalog no. 12191) at a final concentration of 2.05 µg/µl.

Two probes for use in library screening were generated by the polymerase chain

15 reaction (PCR) using 100 ng of library L401 as template. Standard PCR techniques were used. A reaction mixture contained 10 pmol of each oligo primer; 0.2 mM each dATP, dTTP, dCTP and dGTP (PE Applied Biosystems catalog no. N808-0260); 1.5 units Expand[™] High Fidelity Taq DNA polymerase and 1x PCR reaction buffer (Roche Molecular Biochemicals, catalog no. 732-641; 10 mM Tris-HCl, 1.5 mM MgCl₂, 50 mM KCl, pH 8.3).

20 The mixture was incubated at 99°C for one minute, followed by 30 cycles of 96°C for 15 seconds, 50°C for 30 seconds, 72°C for 1 minute with a final incubation at 72°C for 7 minutes (MJResearch DNA Engine[®] Tetra PTC-225). The first was generated using oligos

107335: 5'-CAGCGGCCTGGAGGATGC-3' (SEQ ID NO: 722)

107338: 5'-CGGTCCAGTAGAGGTTTCG-3' (SEQ ID NO: 723)

25 which amplify a *NotI*-*SaII* fragment of the *Zmax1* gene. The second was generated using oligos

107341: 5'-CATCAGCCGCGCCTTCATG-3' (SEQ ID NO: 724)

107342: 5'-CCTGCATGTTGGTGAAGTAC-3' (SEQ ID NO: 725)

which amplify a *SacI*-*KpnI* fragment of the *Zmax1* gene. Both PCR products were purified

using the Qiaquick[®] kit and the manufacturer's protocol (Qiagen catalog 28106) then subcloned into the vector pCRII-TOPO[®] (Invitrogen catalog no. K4600) following the manufacturer's protocol. Positive subclones were identified by restriction digestion of purified plasmid DNA (using standard molecular biology techniques) and subsequent DNA
5 sequence analysis (ABI Prism BigDye Terminator Cycle[®] sequencing, catalog no. 4303154, ABI 377 instruments). Probe DNA was isolated by *Eco*RI restriction digestion (New England Biolabs, catalog no. R0101L) of the respective sequence-verified pCRII-TOPO[®] clone. Restriction fragments were size-fractionated by gel electrophoresis on 1% agarose, 0.1 ug/ml ethidium bromide, 1x TAE gels. Insert DNA was excised from the gel and purified
10 using the Clontech NucleoSpin[®] Nucleic Acid Purification Kit (catalog no. K3051-2) following the manufacturer's protocol. The purified DNA fragment (25-50 ng) was labeled with Redivue[®] (α^{32} P)-dCTP (Amersham Pharmacia, catalog no. AA0005) using the Prime-It II[®] Random Primer labeling Kit and protocol (Stratagene, catalog no. 300385). Unincorporated dCTP was removed with Amersham's NICK[®] column and protocol (catalog
15 no. 17-0855-02).

Two rounds of screening library L401 were initiated to isolate fragments of the *Zmax1* (*LRP5*) gene. In the first, forty-three 150 cm LB -100 μ g/ml ampicillin agar plates were plated with primary transformants from L401 at a density of about 3,000-4,000 colonies per plate. This library was screened using the ³²P-labeled probe (*Not*I-*Sa*I fragment) as described
20 above at 500,000-1,000,000 cpm/ml hybridization buffer, using standard molecular biology protocols. From this primary screen, 13 single colonies were identified based on positive hybridization to the *Zmax1* probe. Plasmid DNA, prepared using the QIAprep Spin Miniprep Kit and protocol (Qiagen Inc., catalog no. 27106), was analyzed by restriction digestion and sequence analysis as described above. cDNA clone # 44 was isolated from this screen and
25 sequence verified to contain a partial *Zmax1* clone.

In the second library screen, one hundred-and-four 150 cm LB ampicillin 100 ug/ml agar plates were plated with primary transformants from L401 at a density of 3000-4000 colonies per plate. This library was screened using the ³²P-labeled probe (*Sac*I-*Kpn*I) exactly as previously described. From this primary screen, 48 colonies were identified based on

positive hybridization to the *Zmax1* (*LRP5*) probe. Since these colonies were not single colony isolates, a secondary screen was initiated where each of the 48 primary isolates was plated at a density of approximately 500 colonies per plate. These colonies were then screened exactly as the primaries using the labeled *SacI-KpnI* fragment as probe. Thirty-four
5 of the 48 primary clones resulted in positive hybridization to the *Zmax1* probe and were isolated as single colonies. Plasmid DNA was prepared and analyzed as described above. cDNA isolate #71_2 was isolated from this screen and sequence verified to contain a partial *Zmax1* clone.

In all cases, the sequence of any *Zmax1* (*LRP5*) isolate was compared to a reference
10 sequence (i.e., the sequence of the wild-type *Zmax1* allele from an affected member of the HBM kindred). This analysis was important since DNA polymorphisms had been reported for this gene in the literature. This reference sequence is predicted to encode a polypeptide of Genbank Accession No. AF077820.

The four independent partial clones used to prepare *Zmax1GI_3AS* are as follows:

15 1). Bases 1-1366: A *XbaI-SalI* fragment was obtained from a *Zmax1* cDNA construct, GTC.*Zmax1_13*. GTC.*Zmax1_13* contains a 5075 BP insert containing the entire ORF of *LRP5*. The clone was blunt end cloned in the *EcoRV* site of pSTBlue-1. This clone was generated by fusing a 5' clone derived from screening a bone random primed cDNA library in a pBluescript™ II derivative with a 3' clone derived
20 from a PCR product from a bone dT primed cDNA library in pBluescript™ II. PCR was performed using *LRP5* specific forward primer 5'-GCCCGAAACCTCTACTGGACCGAC-3' (SEQ ID NO: 726) and reverse primer 5'-GCCCACCCCATCACAGTTCA CATT-3' (SEQ ID NO: 727) using DNAzyme polymerase. The resultant 3.7 kb PCR product was cloned into PCR-XL-TOPO. To generate the full length clone the 5' and
25 3' plasmids were transformed into DM1 (dam-) from Gibco/BRL. The 5' plasmid was digested with *XbaI* and the 3' plasmid was digested with *HindIII*. The digested plasmids were filled in with T4 polymerase to generate blunt ends and cut with *BclI*. the 1.7 kb 5' fragment and 3.5 kb 3' fragments were gel purified, ligated together, and cloned in the *EcoRV* site of pSTBlue-1. It provides a short 5'UTR, with coding

sequence beginning at base 100. Furthermore, it carries some additional restriction sites at the 5' multiple cloning site. This fragment also contains a DNA polymorphism relative to the Genbank Accession No. AF077820 sequence at position 558 resulting in an A (AF077820) to a G change; this mutation does not result in an amino acid difference (Pro).

2). Bases 1367-2403: This clone was obtained from a *Zmax1*-gene primed cDNA library made from commercial human liver RNA described above. This fragment is a *SalI*-*BamHI* piece of DNA obtained from isolate #44. The sequence is identical to Genbank Accession No. AF077820.

3). Bases 2404-4013: This *BamHI*-*BssHII* fragment was obtained from isolate #71-2 from the same library as described above. At position 3456, there is a DNA polymorphism resulting in a G (AF077820) to an A. This nucleotide difference does not change the encoded amino acid (Val).

4). Bases 4014-5278: This *BssHII*-*NotI* fragment came from an internal clone, L236B_P0049E08. It was obtained from an oligo-dT primed HeLa cell cDNA library. The stop codon occurs at base 4947. The clone contains 331 bp of 3' UTR sequence, including a 120 bp poly-A tail followed by the *NotI* site. The 3' *NotI* site used in this subcloning step is a result of an added linker that was introduced at the end of the poly-A tail during library construction. A DNA polymorphism is present at base 4515 resulting in a G (AF077820) to C change that is silent at the amino acid level (Leu).

To generate the 5' section of the *Zmax1* (*LRP5*) gene, the *XbaI*-*SalI* fragment and *SalI*-*BamHI* fragment were ligated into the *XbaI*-*BamHI* sites of pBluescript® (Stratagene). The 3' section of the gene was obtained by ligating the 1.61 kb *BamHI*-*BssHII* fragment from *Zmax1* isolate #71_2 to the 1.26 kb *BssHII*-*NotI* fragment from L236B_P0049E08. These two fragments were ligated into the *BamHI*-*NotI* sites of pBluescript®. The full length *Zmax1* cDNA was engineered into the *XbaI*-*NotI* sites of the vector pCMVSPORT6.0 (Life Technologies) by ligation of the *XbaI*-*BamHI* 5' section and the *BamHI*-*NotI* 3' section. The resulting plasmid, *ZmaxGI_3AS*, contains an insert of 5278 bp from the *XbaI* site to the *NotI*

site of the vector's multiple cloning site. This clone is in the antisense orientation with respect to the CMV promoter present in the vector. *Zmax1* coding sequence begins at base 100 and ends at base 4947, followed by 331 bp of 3'-primed UTR sequence including a 120 bp poly-A tail. This full length cDNA contains three DNA polymorphisms from the reference sequence (GenBank Accession No. AF077820) that do not alter the predicted amino acid sequence. These polymorphisms are at position 558 resulting in an A to G that maintains the proline residue; at position 3456 resulting in a G to A that maintains the valine residue; and, at position 4515 resulting in a G to C that maintains the leucine residue.

The sequence of *Zmax1GI_3AS* (Figure 25) also contains a DNA polymorphism relative to SEQ ID NO. 1 at base 4088 resulting in a C (SEQ ID NO: 1 and Genbank Accession No: AB017498) to T change that results in an amino acid change at position 1330 of alanine to valine. This is consistent with the sequence determined in the wild-type allele from an affected member of the HBM kindred as well as with the published sequence of Genbank Accession No. AF077820. *Zmax1GI_3AS* also has 29 additional bases at the 5' end relative to SEQ ID NO: 1, as well as 129 bases at the 3' end consisting of an extra G, 120 bases of poly-A tract, and the *NotI* site.

21.2.2 Creation of the HBM mutation G171V

The HBM mutation that results in a predicted amino acid change from glycine to valine at amino acid 171 was introduced into the full length human *Zmax1* (*LRP5*) cDNA (plasmid *Zmax1GI_3AS*) using PCR to change the G at position 611 to a T. Introduction of the HBM mutation was done using oligos 107335: (5'-CAGCGGCCTGGA GGATGC-3'; SEQ ID NO: 728) and 49513: (5'-CGGGTACATGTACTGGACAGC TGATTAGC-3'; SEQ ID NO: 729), which flank the endogenous *NotI* site of the *Zmax1* (*LRP5*) gene. This method creates a new *PvuII* site at the 3' end of the PCR product. A second PCR reaction was completed using oligos which introduce a *ScaI* site at the 5' end of the product and contains the endogenous *SalI* site of *Zmax1* in the 3'-primed end. PCR products were purified using the QiaQuick® procedure (Qiagen Inc.); subcloned into the vector pCRII-TOPO (Invitrogen) as described above. Plasmid DNA was purified from single bacterial colonies and analyzed

by restriction digest and subsequent sequence analysis, all as described above. The sequence-verified pCRII-TOPO clones were restriction digested with *NotI*-*PvuII* and *ScaI*-*SaII*, respectively. The resulting DNA fragments were size fractionated and purified as described above. These two fragments were then subcloned into the vector, pBluescript® that had been prepared by *NotI*-*SaII* digestion. Both *PvuII* and *ScaI* produce blunt ends when used to digest double stranded nucleic acids. Thus, the resulting ligated fragment fails to recreate either the *PvuII* or *ScaI* site and contains only the consensus *Zmax1* sequence, with the exception of the newly introduced HBM mutation. To introduce the mutation into the full length *Zmax1* gene, this resulting plasmid was digested with *MscI* and *SaII*, while the 5' region of *Zmax1* was obtained by *XbaI*-*MscI* digestion of *Zmax1* plasmid GTC.*Zmax1*_13. These two fragments were ligated together into *XbaI*-*SaII* digested pBluescript®, in effect creating a similar 1.366 kb *XbaI*-*SaII* fragment. The only difference being that this construct contains the HBM mutation described above. The full length HBM cDNA then was assembled into pCMVSPORT6.0 exactly as described above for the *Zmax1* gene, with the substitution of this newly created *XbaI*-*SaII* fragment containing the HBM mutation. The entire cDNA insert was verified by DNA sequence analysis and the introduction of the HBM mutation was confirmed.

The resulting plasmid, HBMGI_2AS (Fig. 24), contains an insert of 5,278 bp from the *XbaI* site to the *NotI* site of the vector's multiple cloning site. This clone is in the antisense orientation with respect to the CMV promoter in the vector. HBM coding sequence begins at base 100 and ends at base 4947, followed by 331 bp of 3'-primed UTR sequence which includes a 120 bp poly-A tail. This full length cDNA contains three DNA polymorphisms from the reference sequence, which do not alter the amino acid sequence. These polymorphisms occur at position 558, resulting in an A to G change that maintains the proline residue; at position 3456 resulting in a G to A change that maintains the valine residue; and, at position 4515 resulting in a G to C change that maintains the leucine residue. Additionally, the HBM mutation is present at position 611 (G in *Zmax1* to T in HBM) which results in a predicted amino acid change of glycine to valine at amino acid position 171, as found in affected members of the HBM kindred. This insert sequence was used to generate

the construct used for HBM over-expressing transgenic mice.

21.3 Transgene preparation

The examples provided herein are illustrations of how transgenic animals can be prepared. Additional transgenic animals can be prepared as would be known in the art. See, for example, Glenn Monastersley *et al.*, ed, Strategies in Transgenic Animal Studies (Amer. Soc. Microbiology 1995) and the references cited therein.

21.3.1 CMVbActin promoter-HBM cDNA (HBMMCBA)

To prepare the CMVbactin-HBM construct, pCX-EGFP, a plasmid containing the chimeric CMVbactin promoter, was purified as a 4778 bp *EcoRI* fragment. Subsequently, the HBM cDNA was excised from HBMGI_2AS as a 4994 bp *XbaI/DraI* fragment, treated with Klenow fragment of DNA polymerase, ligated to *EcoRI* linkers, and digested with *EcoRI*. This fragment was then inserted into the *EcoRI* site of pCX-EGFP. A *SpeI/HindIII* 7265 bp CMVbactin-HBM fragment was purified for microinjection into mouse embryos.

21.3.2 Type I collagen promoter-HBM cDNA (HBMMTIC)

The rat type I collagen promoter-HBM construct was created by first replacing the pBS(SK-) (Stratagene) polylinker with another polylinker (i.e., comprising *KpnI* - *SpeI* - *HindIII* - *BglII* - *NdeI* - *SaII* - *SmaI* - *EcoRI* - *PstI* - *BamHI* - *XbaI* - *ScaI* - *NcoI* - *ClaI* - *NotI* - *SacII* - *SacI*), that is referred to as BS(SK-)A/D. The SV40 splice and poly (A)_n *XbaI*-*NcoI* region (750 bp) from pcDNA I (Invitrogen, Inc.) was directionally cloned into BS(SK-)A/D. Next, a 4994 bp *EcoRI* HBM cDNA fragment (above) was cloned into the *EcoRI* site. A 3640 bp *XbaI*, type I collagen promoter fragment was subcloned into the *XbaI* site of BS(SK-) (Stratagene). The promoter fragment was then excised from BS(SK-) with *SacII*, blunt-ended with T4 DNA polymerase, digested with *SpeI*, and ligated into the HBM BS(SK-) A/D construct, which was digested with *NdeI*, blunted with T4 DNA polymerase, and digested with *SpeI*. A *SpeI/ClaI* 9435 bp type I collagen-HBM fragment was purified for microinjection into mouse embryos.

21.3.3 CMVbActin promoter-Zmax1 cDNA (Zmax1WTCBA)

The CMVbActin promoter-HBM cDNA construct from above was used to generate the final plasmid. The following three fragments were ligated together: 1) a 6.34 kb *XbaI*-*KpnI* backbone fragment from HBMMCBA; 2) a 0.64 kb *XbaI*-*SapI* fragment from
 5 HBMMCBA containing the 3' end of the bActin promoter and the 5' end of the HBM cDNA; and 3) a 2.8 kb *SapI*-*KpnI* fragment derived from the Zmax1 cDNA that contains the wild-type sequence. A 7.2 kb *SpeI* - *HindIII* CMVbActin - Zmax1 fragment was purified for micro-injection into mouse embryos.

10 21.3.4 Type I collagen promoter-Zmax1 cDNA (Zmax1WTTIC)

The type I collagen - HBM cDNA construct from above was used to generate the final plasmid. The HBMMTIC plasmid was linearized with *HindIII* and cut with either *SalI* to yield a 8.52 kb *HindIII* - *SalI* fragment or *SapI* to yield a 2.98 kb *SapI* - *HindIII* fragment.
 15 A 2.8kb *SapI* - *SalI* fragment from the Zmax1 (LRP5) cDNA containing the wild-type sequence was then ligated to the above two fragments to yield the final plasmid. A 9.4 kb *SpeI* - *ClaI* type I collagen - Zmax1 fragment was purified for micro-injection into mouse embryos.

21.4 Confirmation of transgene expression *in vitro*

20 Plasmid constructs for HBMMCBA, HBMMTIC, Zmax1WTCBA and Zmax1WTTIC were transiently transfected into human osteoblast (HOB) cells to measure mRNA expression as a test for functionality.

21.4.1 Transient Transfections

25 HOB-02-02 cells are a clonal, post-senescent, cell line derived from the HOB-02-C1 cells (Bodine et. al, 1996, *J. Bone Miner. Res.* 11: 806-819). Like the parental cell line, the HOB-02-02 cells express the temperature-sensitive SV40 large T-antigen mutant, *tsA209*. Consequently, these cells proliferate at the permissive temperature of 34°C, but stop dividing at non-permissive temperatures of 37°C or above. Also like the parental cell line, the HOB-

02-02 cells are cultured with Growth Medium (D-MEM/F-12 containing 10% heat inactivated fetal bovine serum, 1% penicillin-streptomycin and 2 mM GlutaMAX-1) at 34°C in a 5% CO₂/95% humidified air incubator (Forma Scientific, Marietta, OH).

For the transient transfections, the HOB-02-02 cells were seeded with Growth Medium at 400,000 cells/well into 6-well plates and incubated overnight at 34°C. The cells were transfected with 0.3 mg/well of either the CMVβActin-HBM expression plasmid, the Type I Collagen-HBM expression plasmid or the corresponding empty vectors using LipofectAMINE 2000 transfection reagent according to the manufacturer's instructions (Life Technologies, Rockville, Md). After a 24 hr incubation at 34°C, the medium was changed, and the cells were incubated for an additional 24 hr at 39°C. At the end of this last incubation, the cells were rinsed with Hank's buffered salt solution. Total cellular RNA was then isolated using TRIzol® according to the manufacturer's instructions (GibcoBRL, Grand Island, NY). The RNA was treated with RNase-free DNase in order to remove contaminating DNA as previously described (Bodine *et al.*, 1997, *J. Cell. Biochem.* 65: 368-387).

21.4.2 TaqMan® assay for mRNA expression

TaqMan® primers and probes were chosen based on human and mouse Zmax1 (LRP5) cDNA sequences. The selected sequences were designed to be gene-specific by analysis of an alignment of human and mouse Zmax1 (LRP5) sequences as illustrated in Fig. 26.

TaqMan® quantitative reverse transcriptase-polymerase chain reaction (RT-PCR) analysis of RNA isolated from human cells was performed as described by the manufacturer (PE Applied Biosystems, Foster City, CA) using the following primers and probe set:
Human Zmax1-1/HBM-1:

Forward Primer: 5'-GTCAGCCTGGAGGAGTTCTCA-3' (SEQ ID NO: 730)
Reverse Primer: 5'-TCACCCTTGGCAATACAGATGT-3' (SEQ ID NO:731)
Probe: 6-FAM-5'-CCCACCCATGTGCCCCGTGACA-3' (SEQ ID NO:732)

Results from the experimental primers/probe set were normalized to human GAPDH

levels using the multiplex protocol with the human GAPDH control kit from PE Applied Biosystems. Species-specific TaqMan® quantitative RT-PCR analysis of RNA isolated from murine cells and tissues was performed as described by the manufacturer (PE Applied Biosystems, Foster City, CA) using the following primers and probes sets:

5 Human Zmax-1/HBM-1:

Forward Primer: 5'-CGTGATTGCCGACGATCTC-3' (SEQ ID NO:733)
 Reverse Primer: 5'-TTCCGGCCGCTAGTCTTGT-3' (SEQ ID NO:734)
 Probe: 6-FAM-5'-CGCACCCGTTTCGGTCTGACGCAGTAC-3' (SEQ
 ID NO:735)

10 Mouse Zmax-1/HBM-1:

Forward Primer: 5'-CTTTCCCCACGAGTATGTTGGT-3' (SEQ ID NO:736)
 Reverse Primer: 5'-AAGGGACCGTGCTGTGAGC-3' (SEQ ID NO:737)
 Probe: 6-FAM-5'-AGCCCCTCATGTGCCTCTCAACTTCATAG-3'
 (SEQ ID NO:738)

15 Results from the experimental primers/probe sets were normalized to 18S ribosomal RNA levels using the multiplex protocol with the 18S ribosomal RNA control kit from PE Applied Biosystems. A summary of these results is presented in Figures 17-20.

20 21.5 Production of transgenic mice

21.5.1 DNA microinjection

Transgene fragments for micro-injection were first purified on 1% agarose gels according to the GELase protocol from Epicentre Technologies. Fragments were then further purified on cesium chloride density gradients and extensively dialyzed against 5 mM Tris (pH 7.4), and 0.1 mM EDTA.

25 Linearized DNA was microinjected into mouse embryos according to standard procedures. DNA was injected into primarily the male pronucleus of fertilized C57BL/6T mouse embryos. Injected embryos (n=20-35) were transferred to the oviduct (unilaterally) of day 0.5 post coitum pseudopregnant Swiss Webster embryo recipients. Offspring were tail-biopsied and genotyped at age 10-14 days.

21.6 Production of gene-targeted transgenic mice

21.6.1 Gene targeting vectors and probes

Two gene-targeting vectors were constructed for modification of the *Zmax1* (*LRP5*) gene in embryonic stem (ES) cells. The two constructs, illustrated in Figure 16, designated as *Zmax1*-KI/KO A&B were designed to generate two types of mutations, a knock-out (KO) of the *Zmax1* gene and a Cre recombinase dependent knock-in (KI) of a nucleotide substitution in order to create a mouse model (i.e., glycine 170 to valine amino acid substitution in mouse *Zmax1*, of the HBM kindred.

Both gene-targeting vectors were constructed using genomic DNA of the mouse genomic DNA BAC clone 473P5 (Genbank Accession No. AZ095413) containing the first five exons of the mouse *Zmax1* gene. This clone was isolated by Research Genetics (Huntsville, AL) from their mouse 129SvJ genomic BAC library using a polymerase chain reaction (PCR) screen for exon 3. A forward and reverse primers of

Forward: 5'-GAGCGGGCAGGGATGGATGG-3' (SEQ ID NO: 739)

Reverse: 5'-AGGTTGGCACGGTGGATGAAGC-3' (SEQ ID NO: 740)

were used to amplify exon 3 by PCR; the following thermal cycling conditions were employed: for thirty cycles, 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. Identity of this clone with mouse *Zmax1* was confirmed by sequencing exon 3 using the BAC clone DNA as template. PCR products were cloned using the pGEM-T-easy T/A cloning kit.

21.6.2 *Zmax1* (*LRP5*) knock-in/knock-out vector

The organization of the genomic BAC clone 473P5 was characterized by Southern blot analysis using subcloned exon 1, exon 2 and exon 3 as probes and by sequencing the region spanning exon 1 through exon 5. Two different constructs were prepared for the *Zmax1* KI/KO targeting. These constructs (A and B) differ only in flanking arms of homology. Construct (A) contains a 6.5 kb *Bst*EII-*Xba*I 5' arm of homology and a 1 kb *Xba*I-*Eco*RI 3' arm of homology; whereas, construct (B) contains a 1 kb 5' arm of homology and a 6.0 kb 3' arm of homology. The constructs were prepared by ligating short and long arms of homology to a LoxP flanked cassette containing the neomycin resistance gene (MC1-

Neo, Stratagene) and a synthetic transcriptional pause sequence (Promega).

Both *Zmax1*-KI/KO- targeting vectors (A and B) were modified to a G-to-T nucleotide substitution, encoding the G170V amino acid substitution, in exon 3. These modifications were introduced into by overlapping PCR mutagenesis using the wild type sequence of the short arm of homology as template. In addition, the 1 kb short arm of the *Zmax1*-KI/KO (B) targeting vector was modified to include a 5' terminal *PmeI* restriction recognition site. The 5' overlapping fragment was made using the following forward primer and reverse mutagenic primer:

Forward: 5'-AAGCTTGTTTAAACTGGGCATGGTGGCACATGGTTGTAAT-3' (SEQ ID NO: 741)

Reverse: 5'-GGGCTTCCACCCAGTCAGTCCAGTACATGTACCT-3' (SEQ ID NO: 742).

The thermal cycling conditions utilized for thirty cycles are 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. The 3' fragment was made using the following forward and reverse primers:

Forward: 5'-CTGACTGGGTGGAAGCACCCCGGATCGAGC-3' (SEQ ID NO: 743)

Reverse (mutagenic): 5'-GAATTCATCGGTACCTGTGCGGCCGCTTCATTG -3' (SEQ ID NO: 744).

The thermal cycling conditions utilized for thirty cycles are 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. The final overlapping PCR used 1 ml each of the 5' fragment and 3' fragment PCR reactions as template and amplification was performed using the forward and reverse primers of the 5' and 3' fragments respectively and the same thermal cycling parameters. The final PCR product was cloned using the pGEM-T-easy T/A cloning kit. The mutagenized exon 3 was excised from *Zmax1*-KI/KO (B) and transferred to *Zmax1*-KI/KO (A) as a 600 bp *BsmBI*-*XbaI* fragment.

Probes for screening for and characterization of *Zmax1*-KI/KO (A) gene targeted ES cell clones are prepared by subcloning restriction fragments of BAC clone 473P5. The 5' outside probe is a 400 bp *Nde*-*BstEII* fragment, and the 3' outside probe is a 500 bp *EcoRI*-*BstXI* fragment.

The outside probes for *Zmax1*-KI/KO (B) are prepared by PCR cloning genomic fragments flanking and immediately adjacent to the targeting vector region of homology. The 5' outside probe used for *Zmax1*-KI/KO (B) is a 498 bp fragment generated using the forward primer of the sequence (5'-TGAGATGTCCTGTCTGTGGC-3'; SEQ ID NO: 745) and a reverse primer of the sequence (5'-TCCTTCCTTCCCTACAGTTG-3'; SEQ ID NO: 746). The thermal cycling conditions utilized with these probes for thirty cycles are: 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. The 3' outside probe is a 600 bp fragment generated using the forward primer of the sequence (5'-CCTAAGGATCTCCTTGTGTCTGTGG-3'; SEQ ID NO: 747) and a reverse primer of the sequence (5'-CTGCAGCAGGTCAGTAGCCTGC-3'; SEQ ID NO: 748). The thermal cycling conditions utilized with these probes for thirty cycles were: 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. Both probes are specific for the *Zmax1* gene in genomic southern analysis. PCR products are cloned using the pGEM-T-easy T/A cloning kit.

A probe for ribonuclease protection analysis of *Zmax1* mRNA structure and transcription levels was prepared by PCR cloning a cDNA fragment containing exon 3 through exon 4. The PCR reaction used a complete cDNA as template, a forward primer of the sequence (5'-TGAGATGTCCTGTCTGTGGC-3'; SEQ ID NO: 749), a reverse primer of the sequence (5'-TCCTTCCTTCCCTACAGTTG-3'; SEQ ID NO: 750) and the following thermal cycling conditions for thirty cycles; 95°C for 0.5 minute, 55°C for 1 minute and 72°C for 1 minute. The PCR product is cloned using the pGEM-T-easy T/A cloning kit.

One skilled in the art could use similar protocols to generate other engineered KI-alleles to produce transgenic animals with an HBM-like phenotype.

21.6.3 Gene targeting in ES cells

For gene targeting, embryonic stem (ES) cells are electroporated with 50 mg of linearized targeting vector and selected in 200 mg/ml G418 for 7-10 days beginning the day after electroporation. G418 resistant clones are picked, expanded and cryopreserved. Resistant clones were screened for homologous recombination by an *EcoRI* genomic Southern restriction fragment length analysis using the 5' outside probe, which detects the

wild type and targeted alleles of Zmax1 as 4 kb and 5 kb fragments, respectively. Gene targeted ES cell clones are thawed, expanded, and characterized by *ScaI* genomic restriction fragment length analysis using the 3' outside probe, which detects the wild type and targeted alleles of Zmax1 as 9 kb and 8 kb fragments, respectively. Gene targeted clones are also characterized by sequence analysis of Zmax1 exon 3 to ensure that the G to T substitution was included in homologous recombination.

21.6.4 Production of gene targeted mice by blastocyst injection

To generate chimeric mice, gene targeted ES cell clones are thawed, expanded and 9-14 ES cells injected into the blastocoel of 3.5 post coitum (p.c.) host C57BL/6 blastocysts. Injected blastocysts (12-17) are then transferred unilaterally into the uterus of 2.5 p.c. pseudopregnant Swiss Webster embryo recipients and allowed to develop to term. Chimeric males are back-crossed to 129SvEv females and tested for transmission of the targeted allele by PCR genotyping with primers specific to the neomycin resistance gene.

21.6.5 In vitro deletion of the Neomycin resistance cassette via Cre recombinase

To generate Zmax1 (LRP5) KI mice from the Zmax1 KI/KO mice the Neomycin resistance (KO) cassette was deleted by micro injection of a Cre expressing plasmid (2 mg/ml) into the male pronucleus of Zmax1 KI/KO pre-fusion zygotes. Deletion of the KO cassette was confirmed by PCR analysis of the cassette insertion site.

21.6.6 Genotyping transgenic mice

Genomic DNA was isolated from mouse tail snips by digestion in 500 ul buffer containing 50 mM Tris-HCl, (pH 7.2), 50 mM EDTA, (pH 8.0), 0.5% SDS and 0.8 mg/ml proteinase K. Samples are incubated at 55°C with shaking overnight. A 10 µl aliquot was heat-inactivated at 99°C for 5 minutes and diluted 1:20 in water. For PCR, 1 µl of the diluted DNA was amplified under the following conditions: Denature: 96°C for 4.5 min; 45 cycles: 96°C for 30 sec; 63°C for 1 min; 72°C for 1 min; Extension: 72°C for 5 min; 4°C hold.

The following primer sets are used for genotyping:

HBMMCBA:

5' primers: 296 bp fragment (SEQ ID NOS: 751-752 respectively)

Forward: 5'-GCT TCT GGC GTG TGA CCG GCG-3'

Reverse: 5'-GCC GCACAG CGC CAG CAG CAG C-3'

5 3' primers: 400 bp fragment (SEQ ID NOS: 753-754 respectively)

Forward: 5'-CAC CCA CGC CCC ACA GCC AGT A-3'

Reverse: 5'-ATT TGC CCT CCC ATA TGT CCT TCC-3'

HBMMTIC:

5' primers: 382 bp fragment (SEQ ID NOS: 755-756 respectively)

10 Forward: 5'-TTC CTC CCA GCC CTC CTC CAT CAG-3'

Reverse: 5'-GCC GCA CAG CGC CAG CAG CAG C-3'

3' primers: 524 bp fragment (SEQ ID NOS: 757-758 respectively)

Forward: 5'-GAA TGG CGC CCC CGA CGA C-3'

Reverse: 5'-GCT CCC ATT CAT CAG TTC CAT AGG-3'

15

21.6.7 Confirmation of genotype by Southern analysis

Mouse genomic DNA was digested with *EcoRI* and probed with a 1.0 kb *Sall-BamHI* restriction fragment from the *Zmax1* cDNA. The probe hybridizes to a 5 kb fragment in transgene positive animals.

20

21.7 Phenotyping

Both *in vivo* and *ex vivo* assays are used to evaluate the phenotype in transgenic mice. Two strains of wild-type mice, namely C57BL/6 and 129 SvEv, are studied to provide control data for phenotypic evaluation in transgenic and gene-targeted mice. In addition, non-transgenic littermate animals are used as controls.

25

21.7.1 In vivo analysis

Spinal bone mineral content (BMC) and bone mineral density (BMD) measurements performed at Creighton University (Omaha, Nebraska) were made by DXA using a Norland

Instruments densitometer (Norland XR2600 Densitometer, Dual Energy X-ray Absorptiometry, DXA). Spinal BMC and BMD at other locations used the machinery available. There are estimated to be 800 DXA machines currently operating in the U.S. Most larger cities have offices or imaging centers which have DXA capabilities, usually a Lunar or Hologic machine. Each location that provided spine BMC and BMD data included copies of the printouts from their machines to provide verification that the regions of interest for measurement of BMD have been chosen appropriately. Complete clinical histories and skeletal radiographs were obtained.

The HBM phenotype (and HBM like phenotype which is also included when discussing the HBM phenotype) in human and animal subjects, preferably humans, can be described using criteria such as: very high spinal BMD; a clinical history devoid of any known high bone mass syndrome; and skeletal radiographs showing a normal shape of the appendicular skeleton.

pDXA: Wild-type and transgenic mice are anesthetized, weighed and whole-body X-ray scans of the skeleton generated using the LUNAR small animal PIXImus device. Scans are begun when the mice are weaned (i.e., at 3 weeks of age) and repeated at 2 week intervals. Wild-type animals are scanned at 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 27, and 29 weeks. Scanning of transgenic animals would be performed for periods up to 17 weeks. Scans are analysed for BMD (bone mineral density), BMC (bone mineral content), TTM (total tissue mass), and % fat for various body regions.

Faxitron radiographs: Following pDXA scanning of anesthetized animals, an additional X-ray was taken using a Faxitron device allowing measurement of bone size.

Calcein labeling: Animals are dosed with 15 mg/kg calcein intraperitoneally on two consecutive occasions. The first dose was given 9 days before euthanasia and the second given 2 days before euthanasia allowing measurement of bone formation rate.

21.7.2 *Ex vivo* analysis

RNA isolation: Total RNA was isolated from tibia and other tissues using TRIzol® to determine mRNA expression.

pQCT: The right femur was cleaned of soft tissue and stored in 70% ethanol for determination of total and trabecular density of the distal metaphysis and cortical density of the mid-shaft.

MicroCT: The right femur was used to determine trabecular indices of the distal metaphysis.

Histology: The right femur was used to determine bone area and static and dynamic parameters of the distal metaphysis.

Bending strength: The left femur was cleaned of soft tissue and stored at -20°C prior to analysis of 3-point bending strength of the mid-shaft.

Compressive strength of vertebra: The entire spine was removed from T10-L6-7. Soft tissue was left on and the spine frozen at -20°C until analysis. Compressive strength was measured at the L5 vertebra.

Serum: Animals are euthanized and serum prepared from blood to measure total cholesterol, triglycerides, osteocalcin and other biochemical markers.

Lysis: Examples include immunocytochemistry, such as *in situ* hybridization of osteogenic markers and TUNEL staining of cells undergoing apoptosis.

21.8 Results

21.8.1 Confirmation of expression from transgenic plasmid constructs

The HBM (HBMMCBA and HBMMTIC) and wild-type (Zmax1WTCBA and Zmax1WTTIC) plasmid constructs were transiently transfected into HOB-02-02 cells, which have a very low endogenous level of Zmax1 (LRP5) expression. Two days after transfection, RNA was isolated and TaqMan® quantitative RT-PCR was performed to determine the mRNA levels of Zmax1/HBM in the cells. To control for contaminating plasmid DNA, PCR was performed with or without the prior RT step, in the absence of the RT step, only very low levels of Zmax1/HBM mRNA were detected. However, with the RT step, a 1000-fold increase in HBM and Zmax1 mRNA was observed in cells transfected with CMVbActin-promoter constructs as compared to those transfected with the CMV-b-Galactosidase control. The type I collagen promoter constructs showed approximately 10-fold increases in HBM

and Zmax1 mRNAs, which is consistent with the weaker nature of this promoter compared to the CMV β Actin promoter. See Fig. 17.

21.8.2 Species specific TaqMan® reagents for HBM/Zmax1 expression

Species specific TaqMan® primer and probe sets for Zmax1/HBM were developed. In a series of experiments using HOB cells and mouse MC-3T3-E1 osteoblastic cells, Zmax1/HBM mRNA was measured in a mouse background, and vice versa. These reagents useful for the detection and quantization of species-specific expression. As demonstrated in Fig. 18, the primer sets are species specific in the mouse and human cell lines. Further, Fig. 19 demonstrates the quantitative measurement of human Zmax1 RNA in a background of mouse RNA. These TaqMan® sets can be used to determine the levels of human or mouse HBM or Zmax1 (or other HBM-like variant) message that are being expressed in the mouse transgenic lines.

The species-specific TaqMan® reagents are novel tools for the characterization of both endogenous Zmax1 mRNA levels and human Zmax1/HBM mRNA levels in the transgenic mouse tissues. These tools have several advantages over other conventional methods, such as Northern hybridization and standard RT-PCR. Some of these advantages are as follows: (1) *specificity, since only a small region (<100 bps) is amplified primers and probes are chosen to sequence regions predicted to have no or minimal cross-reactivity;* (2) *speed, since the procedure is less labor intensive;* (3) *accuracy, since it is truly quantitative;* and (4) *sensitivity, since it requires only small amounts of starting material (i.e., RNA) and the signal-to-noise ratio is high.* These advantages are especially important for analyzing mRNA levels in bone, because it is difficult to obtain large amounts of RNA from bone. Thus, the primer sets developed for TaqMan® analysis of a HBM and Zmax1 expression are important embodiments of the present invention. One skilled in the art will recognize that the primers described here are preferred embodiments; modifications such as extension or truncation of a primer or base substitution are encompassed by the present invention so long as the resultant nucleic acid continues to perform substantially the same function.

21.8.3 HBM expression in transgenic mice

Eight transgenic founder animals were produced for the CMV β Actin (HBMMCBA) construct and a breeding program initiated to establish lines. Expression of mRNA determined by TaqMan® analysis, shown in Fig. 20, showed variable levels of bone expression in 4 lines. In tibia, expression levels (relative to endogenous Zmax1 (LRP5) in HOB-03-C5 cells) showed the following range: line 18 (x10-11 fold); line 2 (x7-10 fold); line 13 (x1-2 fold) and line 28 (x1 fold). Expression was also detected in other tissues as expected based on the known activity of the promoter. For lines 2 and 13, the highest levels of HBM expression were found in the heart. A TaqMan® genotyping assay will screen for potential homozygous animals.

Six transgenic founder animals were produced for the type I collagen (HBMMTIC) construct, and a breeding program initiated to establish lines. Expression of mRNA was found in two lines initially tested. In line 19, expression was 7-8 fold and 19-20 fold greater than Zmax1 in HOB-03-C5 cells in tibia and femur, respectively. In line 35, a low level of expression was detected in tibia and femur.

21.9 In vivo pDXA in HBM transgenic mice

21.9.1 HBMMCBA construct

Initial analysis of limited numbers of mice, illustrated in Fig. 21 (A-C), at the 5 week and 9 week time-points showed that one line tested to date had greater BMD values compared to control. At 5 weeks, HBMMCBA line 2 (n=11) BMD in femur, spine and total body was 21%, 24% and 10% greater respectively, than wild-type control. At 9 weeks (n=3), these increases in BMD amounted to 19%, 32% and 12%, respectively.

21.9.2 HBMMTIC construct

Analysis of limited numbers of mice, illustrated in Fig. 21 (D-F), at the 5 week and 9 week time points showed that two lines tested to date had greater BMD values as compared to control. At 5 weeks, HBMMTIC line 19 (n=5) BMD in femur, spine and total body was 63%, 70% and 41% greater respectively than the wild-type control. At 9 weeks (n=2), these

increases in BMD amounted to 52%, 64% and 37% respectively.

At 5 weeks, HBMMTIC line 35 (n=1) BMD in femur, spine and total body was 4%, 47% and 6% greater, respectively than the wild-type control. At 9 weeks (n=3), these increases in BMD amounted to 32%, 43% and 18% respectively.

5 These results indicate that three lines analyzed to date show evidence for a high bone mass phenotype. There appeared to be no obvious correlation between levels of mRNA expression and BMD phenotype from the limited numbers of lines and animals studied to date. For example, HBMMCBA line 2 and HBMMTIC line 19 have similar levels of HBM mRNA in tibia, but the phenotype is more evident in line 19. Also, HBMMTIC line 35
10 shows a very low level of expression when compared to HBMMCBA line 2, but HBMMTIC line 35 appears to have a stronger phenotype. These observations point to possible differences in cellular expression that may impact the phenotype.

Overall, the BMD results from the transgenic mice show similarities in magnitude to the phenotype observed in the HBM affected kindred (Johnson *et al.*, 1997, *Am. J. Hum. Genetics*, 60:1326-1332). For example, spinal BMD measured in affected individuals is
15 approximately 34-63% greater than non-affected family members. The data for spinal BMD from the transgenic animals ranges from ~30-70% greater than normal at 9 weeks of age.

21.9.3 *Ex vivo* analysis of transgenic mice

20 In order to further examine increases in bone density that were detected in select transgenic lines through monitoring of the animals by non-invasive bone imaging, necropsies were performed on animals of these lines at 5 and 9 weeks of age for direct bone densitometric and histologic analysis. The left femur was isolated, cleaned and positioned in an XCT Research peripheral Quantitative Computed Tomograph (pQCT; Stradtec
25 Medizintechnik, Pforzheim, Germany). The distal end of the femur was located and pQCT scanning was initiated 2.5 mm proximal from this point for total and trabecular measurements. The pQCT scan for cortical measurements was initiated 3.5 mm proximal from the first scan (i.e., 6 mm proximal from the distal end). The pQCT scans were 0.5 mm thick, had a voxel (i.e., three dimensional pixel) size of 0.07 mm, and consisted of 360

projections through the slice. After the scans were completed, the images were displayed on the monitor and a region of interest, including the entire femur for each scan, was outlined. The soft tissue was automatically removed using an iterative algorithm, and the density of the remaining bone (total density) in the first slice was determined. The outer 55% of the bone was then peeled away in a concentric spiral and the density of the remaining bone (trabecular density) of the first slice was reported in mg/cm^3 . In the second slice, the boundary between cortical and trabecular bone was determined using an iterative algorithm, and the density of the cortical bone was determined.

Analysis of Line 2 F1 CMV β Actin-HBM 5 week old transgenic animals revealed that total density, trabecular density and cortical density, were 20%, 37% and 4% higher, respectively, in the transgenic male mice versus the non-transgenic males. In 5 week old Line 19 type I collagen-HBM transgenic males, an even more dramatic increase in bone density over their non-transgenic littermates was evident. Total density, trabecular density and cortical density were 53%, 104 % and 5% higher, respectively. In the Line 19 animals, the phenotype was found to be maintained at 9 weeks of age with total, trabecular and cortical bone density being 44%, 101% and 6% higher than the non-transgenic littermates. At 17 weeks, total and trabecular density were increased 46%, 202%, respectively. The effects on the total trabecular parameters in line 19 at all three time points were statistically significantly higher ($p < 0.001$). A somewhat different pattern of bone phenotypic expression was evident from males of type I collagen-HBM transgenic Line 35. At 5 weeks of age total, trabecular and cortical density were only marginally higher (7%, 4% and 4%, respectively). However, at 9 weeks of age a clear and statistically significant ($P < 0.05$, 0.005, and 0.001 respectively) increase in these parameters became evident (i.e., 25%, 37% and 4%, respectively). The occurrence of different patterns of age-related expression of the phenotype is not unexpected, particularly with the “bone specific” type I collagen transgene, which is influenced by stage of bone cell differentiation. Both Line 2 and Line 19 animals at 5 weeks of age express comparable levels of HBM mRNA in tibia samples, and these levels are significantly greater (>7 -8-fold) than other lines that show no apparent bone phenotype at this age. Line 19, which is driven by the type I collagen promoter, unlike line 2, shows very low

expression of the transgene in tissues other than bone. At 5 weeks of age, Line 35 animals show low level expression in bone and none in other tissues. Immunohistochemistry of calvarial bone sections using an HBM/Zmax1 specific antibody reveals much more intense staining in bone cells of transgenic animals from Lines 2 and 19 at 5 weeks of age and from Line 35 at 9 weeks of age versus their non-transgenic littermates.

The findings revealed by pQCT analysis were further examined under greater resolution using μ CT instrumentation (Scanco). The femur was positioned such that the region being imaged includes the distal end of the femur extending approximately 4 mm proximally with the view being perpendicular to the axis of the articulating cartilage. The reference line for beginning the μ CT measurement was placed to minimally overlap the growth plate and extends proximally for 200 scan slices (9 mm thickness). After completing the μ CT measurement, the first slice in which the condyles have fully merged was identified. A region of interest was outlined to include a maximum amount of the trabecular space, while excluding the cortex. For the first thirty slices, regions of interest were drawn every five slices and merged. For the remaining 105 slices, regions of interest were drawn every 10-20 slices. The more regular the trabecular space, the less frequently a region of interest needed to be drawn. Each region of interest was merged with its predecessor after it was drawn. After regions of interest had been established for all 135 slices, three dimensional evaluation was performed using a threshold setting of 350.

The increased bone densities identified by pQCT were confirmed and extended by μ CT to include elements of bone architecture. In the Line 2 transgenic animals, μ CT bone volume/total volume, connectivity density and trabecular thickness were 50%, 83% and 12% higher, respectively. Both the connectivity density and trabecular thickness indices suggest that the increased density is also associated with increased structural strength. Bone surface/bone volume was lower by 17% in the transgenic males, which may suggest that there may be fewer resorptive surfaces and pits. The trabecular bone response was further confirmed by histological evaluation of non-decalcified, Goldner's stained sections, which revealed 36% greater bone mineral area in the distal femoral metaphysis of the transgenic males. Dynamic histomorphometric analysis revealed that a substantial increase in bone mineral apposition

rate (+100%), as determined by calcein double labeling, may be partially responsible for the increased bone in the transgenics. The dramatic effects evident by pQCT on trabecular bone in Line 19 were supported by μ CT evaluation where bone volume/total volume, trabecular number, trabecular thickness and connectivity density were found to be 130%, 45%, 30% and 121% higher, respectively, in the transgenic males. As seen in Line 2, the bone surface/ bone volume was lower in the transgenic males (-36%). All of these effects were statistically significant with $p < 0.01$. The bone phenotype seen at 5 weeks of age in Line 19 was maintained in 9 week-old animals where bone volume/total volume, trabecular number and connectivity density were 125%, 38% and 110% higher than in the non-transgenic littermates.

μ CT analysis of the Line 35 transgenics revealed a somewhat different pattern than the other two lines. In contrast to only modestly increased density indicated by pQCT in 5 week-old females from Line 35, a statistically significant effect ($p < 0.01$) was seen with mCT, which has greater image resolution and encompasses a larger volumetric sample. Bone volume/total volume, trabecular thickness and connectivity density were 35%, 9% and 27% higher. A similar result was seen in 5 week-old males from Line 35 where bone volume/total volume and connectivity density increases of 37% and 45%, respectively, were evident by μ CT analysis, where only slight increases were revealed by pQCT. The differences between the Line 35 transgenic males and their non-transgenic littermates appeared to increase with age such that statistically significant increases in total density (25%) and trabecular density (37%) were evident by pQCT at 9 weeks of age. The μ CT results support an age-related divergence in bone phenotype in this line and show that differences between transgenic and non-transgenic animals, in terms of bone volume/total volume and connectivity density, had nearly doubled those seen at 5 weeks to 70% and 83%, respectively. The bone volume increases seen in the transgenic animals is in agreement with a significant increase in this parameter that was detected in a bone biopsy sample from an adult male affected member of the HBM kindred. The other parameters that were found to be affected in the transgenic lines may reflect changes that lead to an increased peak/adult bone mass, which in this strain of mice occurs between the ages of 17-20 weeks

The bone density and bone architectural changes seen in the over-expressing transgenic lines would suggest potentially greater bio-mechanical strength. This was tested directly by evaluating 3-point bending strength of femurs from 5 week old Line 19 males. The femora were cleaned of soft tissue and the femoral length measured using a digital caliper. Periosteal and endosteal circumferences, as well as cortical thickness, were measured 6 mm from the distal end of the bone using pQCT. The femur was then placed on a fixture so that the center of mid-shaft was at an equal distance from fixed supports located 5 mm apart. The cross bar of an Instron 5543 load device was placed over the mid-shaft and a force applied at a speed of 1 mm/minute until fracture occurred. A force vs. displacement curve was generated and peak load determined using Instron Merlin software. There was a 75% increase ($p < 0.01$) in strength that appears to be due to an increase in periosteal circumference leading to an increase in cortical thickness. Thus, it appears that the changes in bone density and bone geometry, as seen in the HBM transgenic animals, do translate into increases in biomechanical strength.

In view of the association of HBM/Zmax1 within the class of LDL related receptor proteins, it was of interest to determine whether the mutation might affect lipid profiles. Indeed, lipid studies in the HBM kindred (i.e., 8 affected and 7 unaffected members) have revealed that triglyceride and VLDL levels are statistically lower in the affected members. Serum samples from the transgenic lines were analyzed on a Hitachi 911 instrument using Boehringer Mannheim (for Cholesterol) and Roche (for triglycerides) reagents. The cholesterol was measured via *o*-quinone imine dye (which is formed following enzymatic reactions with cholesterol) photometrically at 505 nm at 37°C. Enzymatic methods for triglyceride measurements are based on determination of the glycerol part of triglyceride after hydrolysis of triglycerides and fatty acids. The end dye product of enzymatic reaction was measured at 505 nm. In 5 week old male Line 2 transgenics, although serum cholesterol was only slightly reduced, serum triglyceride levels were reduced by 26% in the transgenics versus their non-transgenic littermates. In a limited sample of Line 2 animals at 9 weeks of age, triglyceride levels remained 20% lower. Similarly, at 5 weeks of age triglyceride levels in male transgenics from Line 19 were 32% lower. In contrast, at 5 weeks of age both male

and female transgenics of Line 35 did not have lower triglyceride levels. The fact that the 5 week old Line 35 animals did have statistically greater bone volume/total volume suggests that the lipid change may not be directly related to the skeletal phenotype. This would appear to be supported by the fact that the Line 35 animals at 9 weeks of age had only slightly
5 reduced triglyceride levels (11%) but exhibited substantially higher bone density than at 5 weeks of age. Due to the different levels and sites of expression of the transgene in these lines we can not rule out the possibility that serum lipid levels could serve as a surrogate marker for agents favorably affecting a bone phenotype through HBM/Zmax1.

These and other transgenic lines based on HBM or HBM-like genes will serve as
10 valuable models for exploring the nature of bone homeostasis. Bone density in all species accommodates to its customary loading conditions. In the HBM kindred and in the transgenic lines, the sensor/effector systems of the skeleton appear to perceive greater load signals resulting in greater bone density. Experimental models have been established showing that increased bone loading can lead to increased bone density and that unloading or
15 disuse leads to a loss of bone density. Evaluating the histological, biochemical and genetic responses of the skeleton of the transgenic animals in these experimental paradigms will yield much information regarding the sensor/effector system responsible for bone homeostasis. The application of the transgenic animals in other established models of altered bone turnover, including but not limited to steroid deficiency-induced osteopenia and aging-related
20 osteopenia will provide further insight into the role of Zmax1 in bone homeostasis and the nature of the favorable changes induced by the HBM mutation.

21.9.4 *HBM* gene-targeting

The Zmax1 KI/KO gene-targeting vector is electroporated into 129 SvEv, C57BL/6
25 ES and 129 ES cells. Restriction fragment length analysis of genomic DNA and sequencing of PCR amplified fragments can be used to identify gene targeted clones. The knock-in version of the gene-targeting vector allows for the introduction of the HBM mutation into the endogenous Zmax (LRP5) genomic locus with minimal impact on the mouse genome. It permits the production of the HBM protein in a more natural environment, i.e. not in an over-

expression model such as the transgenic mice or transfected cell lines. The knock-out version of the gene-targeting vector was engineered to contain a transcriptional stop sequence that has the potential to result in loss of one functional *LRP5* allele. Breeding heterozygous animals with this mutation would lead to the production of embryos homozygous for the null allele.

5 In a different design of the gene-targeting vector, lox P sites could be positioned in such a way that would facilitate production of a conditional knock-out of the endogenous *LRP5* gene. In the presence of Cre recombinase, a critical region of the *LRP5* gene would be deleted in between the lox P sites, thus resulting in the potential loss of one functional allele.

Animal breeding would then be used to create homozygotes with a null allele. Other
10 recombinase enzyme systems, such as flip recombinase in combination with cognate *frt* sites, could be used to create the deletion. The recombinase could be administered in a number of ways as described earlier, including plasmid injection into embryos and use of transgenic animals expressing Cre. The promoter used to drive expression of Cre could be chosen in a manner that would result in ubiquitous or tissue-specific deletion of the *LRP5* gene thus
15 resulting in a conditional knockout. In a further embodiment expression of the Cre enzyme itself could be made conditional using inducible systems such as GeneSwitch and Tetracycline paradigms.

21.9.5 Uses of transgenic animals and cells

20 The transgenic animals and cells of the present invention are useful tools in methods for identifying surrogate markers for the HBM phenotype. The surrogate markers provided by the present invention are also useful tools for the assessment and screening of prospective treatments. Individuals carrying the *HBM* gene have elevated bone mass. The *HBM* gene causes this phenotype by altering the activities, levels, expression patterns, and modification
25 states of other molecules involved in bone development. Using a variety of established techniques, it is possible to identify molecules, preferably proteins or mRNAs, whose activities, levels, expression patterns, and modification states are different between systems containing the *Zmax1* gene and systems containing the *HBM* gene. Such systems can be, for example, cell-free extracts, cells, tissues or living organisms, such as mice or humans. For a

mutant form of *Zmax1*, a complete deletion of *Zmax1*, mutations lacking the extracellular or intracellular portion of the protein, or any other mutation in the *Zmax1* gene may be used. It is also possible to use expression of antisense *Zmax1* RNA or oligonucleotides to inhibit production of the *Zmax1* protein. For a mutant form of *HBM*, a complete deletion of *HBM*, mutations lacking the extracellular or intracellular portion of the *HBM* protein, or any other mutation in the *HBM* gene may be used. It is also possible to use expression of antisense *HBM* RNA or oligonucleotides to inhibit production of the *HBM* protein.

Molecules identified by comparison of *Zmax1* systems and *HBM* systems can be used as surrogate markers in pharmaceutical development or in diagnosis of human or animal bone disease. Alternatively, such molecules may be used in treatment of bone disease. *See*, Schena *et al.*, 1995 *Science*, 270: 467-70.

For example, a transgenic mouse carrying the *HBM* gene or *HBM*-like variant in the mouse homologue locus is constructed. A mouse of the genotype *HBM*/+ is viable, healthy and has elevated bone mass. To identify surrogate markers for elevated bone mass, *HBM*/+ (i.e., heterozygous) and isogenic ++ (i.e., wild-type) mice are sacrificed. Bone tissue mRNA is extracted from each animal, and a "gene chip" corresponding to mRNAs expressed in the ++ individual is constructed. mRNA from different tissues is isolated from animals of each genotype, reverse-transcribed, fluorescently labeled, and then hybridized to gene fragments affixed to a solid support. The ratio of fluorescent intensity between the two populations is indicative of the relative abundance of the specific mRNAs in the ++ and *HBM*/+ animals. Alternatively, mRNA may be isolated from wild-type and transgenic animals. cDNA prepared from these samples is transcribed *in vitro* to obtain labeled mRNA for use on custom made or commercially available gene array chips such as are manufactured by Affymetrix. Sets of genes with altered expression as a function of phenotype may be identified by a variety of routine computational analyses. Genes encoding mRNA over- and under-expressed relative to the wild-type control are candidates for genes coordinately regulated by the *HBM* gene or *HBM*-like gene.

One standard procedure for identification of new proteins that are part of the same signaling cascade as an already-discovered protein is as follows. Cells are treated with

radioactive phosphorous, and the already-discovered protein is manipulated to be more or less active. The phosphorylation state of other proteins in the cell is then monitored by polyacrylamide gel electrophoresis and autoradiography, or similar techniques. Levels of activity of the known protein may be manipulated by many methods, including, for example, comparing wild-type mutant proteins using specific inhibitors such as drugs or antibodies, simply adding or not adding a known extracellular protein, or using antisense inhibition of the expression of the known protein (Tamura *et al.*, 1998 *Science*, 280: 1614-7; Meng, 1998 *EMBO J.*, 17: 4391-403; Cooper *et al.*, 1982 *Cell* 1: 263-73).

In another example, proteins with different levels of phosphorylation are identified in TE85 osteosarcoma cells expressing either a sense or antisense cDNA for Zmax1. TE85 cells normally express high levels of Zmax1 (Dong *et al.*, 1998 *Biochem. & Biophys. Res. Comm.*, 251: 784-90). Cells containing the sense construct express even higher levels of Zmax1, while cells expressing the antisense construct express lower levels. Cells are grown in the presence of ^{32}P , harvested, lysed, and the lysates run on SDS polyacrylamide gels to separate proteins, and the gels subjected to autoradiography (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). Bands that differ in intensity between the sense and antisense cell lines represent phosphoproteins whose phosphorylation state or absolute level varies in response to levels of Zmax1. As an alternative to the ^{32}P -labeling, unlabeled proteins may be separated by SDS-PAGE and subjected to immunoblotting, using the commercially available anti-phosphotyrosine antibody as a probe (Thomas *et al.*, 1995 *Nature*, 376: 267-71). As an alternative to the expression of antisense RNA, transfection with chemically modified antisense oligonucleotides can be used (Woolf *et al.*, 1990 *Nucleic Acids Res.*, 18: 1763-9).

Many bone disorders, such as osteoporosis, have a slow onset and a slow response to treatment. It is therefore useful to develop surrogate markers for bone development and mineralization. Such markers can be useful in developing treatments for bone disorders, and for diagnosing patients who may be at risk for later development of bone disorders. Examples of preferred markers are N- and C-terminal telopeptide markers described, for example, in U.S. Patent Nos. 5,455,179, 5,641,837 and 5,652,112, the disclosures of which

are incorporated by reference herein in their entirety. In the area of HIV disease, CD4 counts and viral load are useful surrogate markers for disease progression (Vlahov *et al.*, 1998 *JAMA*, 279(1): 35-40). There is a need for analogous surrogate markers in the area of bone disease.

5 A surrogate marker can be any characteristic that is easily tested and relatively insensitive to non-specific influences. For example, a surrogate marker can be a molecule such as a protein or mRNA in a tissue or in blood serum. Alternatively, a surrogate marker may be a diagnostic sign such as sensitivity to pain, a reflex response or the like.

10 In yet another example, surrogate markers for elevated bone mass are identified using a pedigree of humans carrying the *HBM* gene or *HBM*-like gene. Blood samples are withdrawn from three individuals that carry the *HBM* gene or *HBM*-like gene, and from three closely related individuals that do not. Proteins in the serum from these individuals are electrophoresed on a two dimensional gel system, in which one dimension separates proteins by size, and another dimension separates proteins by isoelectric point (Epstein *et al.*, 1996
15 *Electrophoresis* 17: 1655-70). Spots corresponding to proteins are identified. A few spots are expected to be present in different amounts or in slightly different positions for the *HBM* individuals compared to their normal relatives. These spots correspond to proteins that are candidate surrogate markers. The identities of the proteins are determined by microsequencing, and antibodies to the proteins can be produced by standard methods for use
20 in diagnostic testing procedures. Diagnostic assays for *HBM* proteins or other candidate surrogate markers include using antibodies described in this invention and a reporter molecule to detect *HBM* in human body fluids, membranes, bones, cells, tissues or extracts thereof. The antibodies can be labeled by joining them covalently or noncovalently with a substance that provides a detectable signal. In many scientific and patent literature, a variety
25 of reporter molecules or labels are described including radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents (U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241). The transgenic or genetically modified animals can also serve in a method for surrogate marker identification.

Using these antibodies, the levels of candidate surrogate markers are measured in

normal individuals and in patients suffering from a bone disorder, such as osteoporosis, osteoporosis pseudoglioma, Engelmann's disease, Ribbing's disease, hyperphosphatasemia, Van Buchem's disease, melorheostosis, osteopetrosis, pycnodysostosis, sclerosteosis, osteopoikilosis, acromegaly, Paget's disease, fibrous dysplasia, tubular stenosis, osteogenesis imperfecta, hypoparathyroidism, pseudohypoparathyroidism, pseudopseudohypoparathyroidism, primary and secondary hyperparathyroidism and associated syndromes, hypercalciuria, medullary carcinoma of the thyroid gland, osteomalacia and other diseases. Techniques for measuring levels of protein in serum in a clinical setting using antibodies are well established. A protein that is consistently present in higher or lower levels in individuals carrying a particular disease or type of disease is a useful surrogate marker.

A surrogate marker can be used in diagnosis of a bone disorder. For example, consider a child that presents to a physician with a high frequency of bone fracture. The underlying cause may be child abuse, inappropriate behavior by the child, or a bone disorder. To rapidly test for a bone disorder, the levels of the surrogate marker protein are measured using the antibody described above.

Levels of modification states of surrogate markers can be measured as indicators of the likely effectiveness of a drug that is being developed. It is especially convenient to use surrogate markers in creating treatments for bone disorders, because alterations in bone development or mineralization may require a long time to be observed. For example, a set of bone mRNAs, termed the "HBM-inducible mRNA set" is found to be over-expressed in HBM/+ mice as compared to +/+ mice, as described above. Expression of this set can be used as a surrogate marker. Specifically, if treatment of +/+ mice with a compound results in over-expression of the HBM-inducible mRNA set, then that compound is considered a promising candidate for further development.

This invention is particularly useful for screening compounds by using the Zmax1 or HBM protein or HBM-like proteins or binding fragments thereof in any of a variety of drug screening techniques.

The Zmax1 or HBM protein or fragment employed in such a test may either be free in solution, affixed to a solid support, or borne on a cell surface. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the protein or fragment, preferably in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, for the formation of complexes between a Zmax1 or HBM protein or fragment and the agent being tested, or examine the degree to which the formation of a complex between a Zmax1 or HBM protein or fragment and a known ligand is interfered with by the agent being tested.

Thus, the present invention provides methods of screening for drugs comprising contacting such an agent with a Zmax1 (LRP5) or HBM protein or fragment thereof and assaying (i) for the presence of a complex between the agent and the Zmax1 or HBM protein or fragment, or (ii) for the presence of a complex between the Zmax1 or HBM protein or fragment and a ligand, by methods well known in the art. In such competitive binding assays the Zmax1 or HBM protein or fragment is typically labeled. Free Zmax1 or HBM protein or fragment is separated from that present in a protein:protein complex, and the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to Zmax1 or HBM or its interference with Zmax1 or HBM: ligand binding, respectively.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the Zmax1 or HBM proteins and is described in detail in WO 84/03564. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with Zmax1 or HBM proteins and washed. Bound Zmax1 or HBM protein is then detected by methods well known in the art. Purified Zmax1 or HBM can be coated directly onto plates for use in the aforementioned drug screening techniques. However, non-neutralizing antibodies to the protein can be used to capture antibodies to immobilize the Zmax1 or HBM protein on the solid phase.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of specifically binding the Zmax1 or HBM protein compete with a test compound for binding to the Zmax1 or HBM protein or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide that

shares one or more antigenic determinants of the Zmax1 or HBM protein.

A further technique for drug screening involves the use of host eukaryotic cell lines or cells (such as described above) that have a nonfunctional Zmax1 or *HBM* gene. These host cell lines or cells are defective at the Zmax1 or HBM protein level. The host cell lines or
5 cells are grown in the presence of drug compound. The rate of growth of the host cells is measured to determine if the compound is capable of regulating the growth of Zmax1 or HBM defective cells.

The goal of rational drug design is to produce structural analogs of biologically active proteins of interest or of small molecules with which they interact (e.g., agonists, antagonists,
10 inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the protein, or which, e.g., enhance or interfere with the function of a protein *in vivo*. See, e.g., Hodgson, 1991 *Bio/Technology*, 9: 19-21. In one approach, one first determines the three-dimensional structure of a protein of interest (e.g., Zmax1 or HBM protein) or, for example, of the Zmax1- or HBM-receptor or ligand complex, by x-ray crystallography, by
15 computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a protein may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson *et al.*, 1990 *Science*, 249: 527-33). In addition, peptides (e.g., Zmax1 or HBM protein) are analyzed by an alanine scan (Wells, 1991 *Methods in*
20 *Enzymol.*, 202: 390-411). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon
25 which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of

peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which have, e.g., improved Zmax1 or HBM protein activity or stability or which act as inhibitors, agonists, antagonists, etc. of Zmax1 or HBM protein activity. By virtue of the availability of cloned Zmax1 or HBM sequences, sufficient
5 amounts of the Zmax1 or HBM protein may be made available to perform such analytical studies as X-ray crystallography. In addition, the knowledge of the Zmax1 or HBM protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to X-ray crystallography.

Identified drug candidates (known as "leads") may be further studied by use of
10 transgenic animals. The transgenic animals of the present invention are useful for creating an animal model of bone density modulation which may be used to test and refine drug leads. The transgenic animals described above represent a single example of the Zmax1 and HBM or HBM-like transgenic animals contemplated herein. One skilled in the art is aware of variations and the considerations which will be routinely applied in modifying the present
15 invention to a specific purpose. Examples of the development of transgenic animal models are given for example in Strategies in Transgenic Animal Science (1995, Monastersky and Robl Eds., Washington, DC: American Society for Microbiology) and references therein which are all incorporated herein by reference in their entirety.

As an example, at least two groups of transgenic animals can be created as described,
20 so that one group expresses HBM and another group expresses Zmax1. These animals can be treated with the candidate drug for some time spanning from a few days to the remainder of the animal's life-span. The animals are monitored for changes in bone mass and/or surrogate markers for the HBM phenotype. The transgenic animals used in such a study may express human HBM protein and Zmax1 protein or the homologous HBM and Zmax1 proteins
25 defined for each species or variants thereof. Expression may be driven by a ubiquitous promoter or a bone specific promoter as would be known. It will be informative to compare groups of animals utilizing different promoters.

The transgenic animals of the method according to the present inventions may also comprise knock-in (KI) and/or knock-out (KO) animals, such as mice, which express HBM,

Zmax1, or neither under the control of the animal's native promoter. Such animals may be created by homologous recombination in ES cells as described above (and elsewhere in the literature of the art such as, for example, U.S. Patent Nos. 6,187,991 and 6,187,992 and references cited therein which are incorporated herein in their entirety). The experimental groups of transgenic animals treated with candidate drugs may be monitored by non-invasive means, by the monitoring of surrogate markers as described above, and/or by *ex vivo* analysis of bones from sacrificed animals at given time-points.

Likewise the effect of such treatments as dietary control (e.g. varying intake of vitamins, minerals, proteins, lipids, etc.), ovariectomy, direct administration of all or part of purified HBM or Zmax1 proteins, administration of antisense nucleotides, antibodies against Zmax1 or gene therapy in adults may be investigated by systematic administration of the treatment to transgenic animals according to the invention. Such treatments may include, administration of estrogens, tamoxifen, raloxifene, (or other selective estrogen modulators, SERMs), vitamin D analogs, calcitonin, cathepsin K inhibitors, statins (e.g. simvastatin, pravastatin, and lovastatin), bis-phosphonates, parathyroid hormone (PTH), bone morphogenetic proteins (BMP) as described in U.S. Patent Nos. 6,190,880 and 5,866,364, and combinations of the above compounds.

In view of the homology between Zmax1 and LRP6 and to the LDL receptor and the further observations of markers for cardiac health being modulated in HBM subjects, it is an aspect of the present invention to use the novel research methods disclosed herein to screen known cardio-protective treatments for bone modulating effects. Thereby, the present invention provides therapeutic methods which are both cardio-protective and which improve bone quality. The models are useful for testing drugs and researching lipid modulation effects related to Zmax1, LRP6, HBM and HBM like proteins and nucleic acids.

The effect of various mutations of *Zmax1*, *HBM* and HBM-like genes may be investigated by creation of additional lines of transgenic animals according to the invention, wherein these animals comprise such mutations. By comparison of direct measures of bone development or surrogate markers, an embodiment of the invention provides a useful research tool for screening gene therapy reagents, candidate drug therapies, and elucidating molecular

mechanisms of bone development modulation. One skilled in the art knows how to use the methods of the present invention to achieve these goals.

The present invention provides a method and useful research tools for testing prospective gene therapies. Transgenic knock-out mice are useful for testing prospective gene therapies. A transgenic knock-out animal such as a mouse is created as described above which does not express endogenous Zmax1 or HBM. A prospective gene therapy, such as intravenous injection of a recombinant replication-defective adenovirus encoding the human HBM protein driven by the CMV β Actin promoter, is administered. Parameters of bone density and/or surrogate markers are monitored over time following therapy (Ishibashi *et al.*, 1993 *J. Clin. Invest.* 92: 883-93). A TaqMan® primer set such as that described above may be used to measure expression of transgenic HBM. One skilled in the art knows alternative methods such as the Northern blot method. Comparison of treated and untreated animals both within and between groups of germ-line transgenic animals, knock-out (null allele) background, and wild-type endogenous Zmax1 background animals provides complementary controls for assessing the relative effectiveness of various modalities of gene therapy.

Uses for the transgenic animals models contemplated herein also include, but are not limited to: (1) sources for generating bone cell cultures from the calvaria of the transgenic animals to study bone cell (e.g., osteoblast and osteoclast) function and number; (2) models for studying the effects of estrogen loss by ovariectomizing (ovx) the transgenic animals; (3) models for testing mechanical loading on the bones and other stress/strength tests; (4) breeding models with which to breed to other genetically modified or naturally occurring mutant animals that display bone abnormalities; (Chipman *et al.*, 1993 *PNAS*, 90: 1701-05; Phillips *et al.*, 2000 *Bone* 27: 219-26; Kajkenova *et al.*, 1997 *J. Bone Min. Res.* 12: 1772-79; Jilka *et al.*, 1996 *J. Clin. Invest.* 97: 1732-40; Takahashi *et al.*, 1994 *Bone and Mineral* 24: 245-55); (5) bone mis/disuse models to test the effects of weight bearing or gravity; (6) models for identifying and screening reagents which may or are known to modulate bone metabolism (e.g., PTH, estrogen, vitamin D analogs, bisphosphonates, statins, leptin, BMP, apoE); (7) models for investigating prospective treatments to improve fracture repair.

The transgenic animal models can be analyzed using, but not limited to, such methods as bone densitometry by pDEXA, pQCT and microCT; histology, molecular marker analysis, apoptosis, cell proliferation, cell cycle, mineralization, serum biochemistry, transcriptional profiling, and the like.

5

22. Methods of Use: Avian and Mammalian Animal Husbandry

The *Zmax1* (LRP5) DNA and *Zmax1* (LRP5) protein and/or the HBM DNA and HBM protein (or LRP6 or an HBM-like nucleic acid or protein as also contemplated herein) can be used for vertebrate and preferably human therapeutic agents and for avian and
10 mammalian veterinary agents, including for livestock breeding. Birds, including, for example, chickens, roosters, hens, turkeys, ostriches, ducks, pheasants and quails, can benefit from the identification of the gene and pathway for high bone mass. In many examples cited in literature (for example, McCoy *et al.*, 1996 *Res. Vet. Sci.*, 60: 185-6), weakened bones due to husbandry conditions cause cage layer fatigue, osteoporosis and high mortality rates.
15 Additional therapeutic agents to treat osteoporosis or other bone disorders in birds can have considerable beneficial effects on avian welfare and the economic conditions of the livestock industry, including, for example, meat and egg production.

**23. Methods of use: Diagnostic assays using *Zmax1*-specific oligonucleotides for
20 detection of genetic alterations affecting bone development**

In cases where an alteration or disease of bone development is suspected to involve an alteration of the *Zmax1*, *LRP6*, *HBM* or *HBM*-like gene, specific oligonucleotides may be constructed and used to assess the level of *Zmax1* mRNA or HBM mRNA, respectively, in bone tissue or in another tissue that affects bone development.

25 For example, to test whether a person has the *HBM* gene, which affects bone density, polymerase chain reaction can be used. Two oligonucleotides are synthesized by standard methods or are obtained from a commercial supplier of custom-made oligonucleotides. The length and base composition are determined by standard criteria using the Oligo 4.0 primer
Picking program (Wojchich Rychlik, 1992) or any suitable alternative. One of the

oligonucleotides is designed so that it will hybridize only to HBM DNA under the PCR conditions used. The other oligonucleotide is designed to hybridize a segment of Zmax1 genomic DNA such that amplification of DNA using these oligonucleotide primers produces a conveniently identified DNA fragment. For example, the pair of primers

5 CCAAGTTCTGAGAAGTCC (SEQ ID NO:32) and AATACCTGAAACCATACCTG (SEQ ID NO:33) will amplify a 530 base pair DNA fragment from a DNA sample when the following conditions are used: step 1 at 95°C for 120 seconds; step 2 at 95°C for 30 seconds; step 3 at 58°C for 30 seconds; step 4 at 72°C for 120 seconds; where steps 2-4 are repeated 35 times. Tissue samples may be obtained from hair follicles, whole blood, or the buccal
10 cavity.

The fragment generated by the above procedure is sequenced by standard techniques. Individuals heterozygous for the *HBM* gene will show an equal amount of G and T at the second position in the codon for glycine 171. Normal or homozygous wild-type individuals will show only G at this position. Similar routine procedures may be used to develop assays
15 for other polymorphisms and variants according to the invention.

Other amplification techniques besides PCR may be used as alternatives, such as ligation-mediated PCR or techniques involving Q-beta replicase (Cahill *et al.*, *Clin. Chem.*, 37(9):1482-5 (1991)). For example, the oligonucleotides 5'-AGCTGCTCGTAGCTGTCTCTCCCTGGATCACGGGTACATGTACTGGACAGACTGGGT-3'(SEQ ID NO:34)
20 and T5'-GAGACGCCCCGGATTGAGCGGGCAGGGATAGCTTATTCCCTGTGCCGCA TTACGGC-3' (SEQ ID NO: 35) can be hybridized to a denatured human DNA sample, treated with a DNA ligase, and then subjected to PCR amplification using the primer oligonucleotides 5'-AGCTGCTCGTAGCTGTCTCTCCCTGGA-3' (SEQ ID NO:36) and 5'-GCCGTAATGCGGCACAGGGAATAAGCT-3' (SEQ ID NO:37). In the first two
25 oligonucleotides, the outer 27 bases are random sequence corresponding to primer binding sites, and the inner 30 bases correspond to sequences in the *Zmax1* gene. The T at the end of the first oligonucleotide corresponds to the *HBM* gene. The first two oligonucleotides are ligated only when hybridized to human DNA carrying the *HBM* gene, which results in the formation of an amplifiable 114 bp DNA fragment.

Products of amplification can be detected by agarose gel electrophoresis, quantitative hybridization, or equivalent techniques for nucleic acid detection known to one skilled in the art of molecular biology (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring, NY (1989)).

5 Other alterations in the *Zmax1* gene or the *HBM* gene (or *LRP6* or HBM-like nucleic acids, which are also contemplated herein) may be diagnosed by the same type of amplification-detection procedures, by using oligonucleotides designed to identify those alterations. These procedures can be used in animals as well as humans to identify alterations in *Zmax1* or HBM that affect bone development.

10 Expression of *Zmax1* or HBM in bone tissue may be accomplished by fusing the cDNA of *Zmax1* or HBM, respectively, to a bone-specific promoter in the context of a vector for genetically engineering vertebrate cells. DNA constructs are introduced into cells by packaging the DNA into virus capsids, by the use of cationic liposomes, electroporation, or by calcium phosphate transfection. Transfected cells, preferably osteoblasts, may be studied
15 in culture or may be introduced into bone tissue in animals by direct injection into bone or by intravenous injection of osteoblasts, followed by incorporation into bone tissue (Ko *et al.*, 1996 *Cancer Res.* 56: 4614-9). For example, the osteocalcin promoter, which is specifically active in osteoblasts, may be used to direct transcription of the *Zmax1* gene or the *HBM* gene. Any of several vectors and transfection methods may be used, such as retroviral vectors,
20 adenovirus vectors, or vectors that are maintained after transfection using cationic liposomes, or other methods and vectors described herein.

Alteration of the level of functional *Zmax1* protein or HBM protein affects the level of bone mineralization. By manipulating levels of functional *Zmax1* protein or HBM protein, it is possible to affect bone development and to increase or decrease levels of bone
25 mineralization. For example, it may be useful to increase bone mineralization in patients with osteoporosis. Alternatively, it may be useful to decrease bone mineralization in patients with osteopetrosis or Paget's disease. Alteration of *Zmax1* levels or HBM levels can also be used as a research tool. Specifically, it is possible to identify proteins, mRNA and other molecules whose level or modification status is altered in response to changes in functional

levels of Zmax1 or HBM. The pathology and pathogenesis of bone disorders is known and described, for example, in Rubin and Farber (Eds.), *Pathology*, 2nd Ed., S.B. Lippincott Co., Philadelphia, PA (1994).

A variety of techniques can be used to alter the levels of functional Zmax1 or HBM. For example, intravenous or intraosseous injection of the extracellular portion of Zmax1 or mutations thereof, or HBM or mutations thereof, will alter the level of Zmax1 activity or HBM activity, respectively, in the body of the treated human, animal or bird. Truncated versions of the Zmax1 protein or HBM protein can also be injected to alter the levels of functional Zmax1 protein or HBM protein, respectively. Certain forms of Zmax1 or HBM enhance the activity of endogenous protein, while other forms are inhibitory.

In a preferred embodiment, the HBM protein is used to treat osteoporosis, fracture, or other bone disorder. In a further preferred embodiment, the extracellular portion of HBM or fragment thereof (e.g., the Dkk binding domain) protein is used. This HBM protein may be optionally modified by the addition of a moiety that causes the protein to adhere to the surface of cells. The protein is prepared in a pharmaceutically acceptable solution and is administered by injection or another method that achieves acceptable pharmacokinetics and distribution.

In a second embodiment of this method, Zmax1, HBM, HBM variant, and/or LRP6 levels are increased or decreased by gene therapy techniques. To increase Zmax1 or HBM levels, osteoblasts or another useful cell type are genetically engineered to express high levels of Zmax1 or HBM as described above. Alternatively, to decrease Zmax1 or HBM levels, antisense constructs that specifically reduce the level of translatable Zmax1 or HBM mRNA can be used. In general, a tissue-nonspecific promoter may be used, such as the CMV promoter or another commercially available promoter found in expression vectors (Wu *et al.*, 1996 *Toxicol. Appl. Pharmacol.* 141: 330-9). In a preferred embodiment, a Zmax1 cDNA or its antisense is transcribed by a bone-specific promoter, such as the osteocalcin or another promoter, to achieve specific expression in bone tissue. In this way, if a Zmax1-expressing DNA construct or HBM-expressing construct is introduced into non-bone tissue, it will not be expressed.

In a third embodiment of this method, antibodies against Zmax1, LRP6, HBM-like or HBM are used to modulate its function. Such antibodies are identified herein.

In a fourth embodiment of this method, drugs that are agonists or antagonists of Zmax1 function or HBM function are used. Such drugs are described herein and optimized according to techniques of medicinal chemistry well known to one skilled in the art of pharmaceutical development.

Zmax1 and HBM interact with several proteins, such as ApoE. Molecules that inhibit the interaction between Zmax1 or HBM and ApoE or another binding partner are expected to alter bone development and mineralization. Such inhibitors may be useful as drugs in the treatment of osteoporosis, osteopetrosis, or other diseases of bone mineralization. Such inhibitors may be low molecular weight compounds, proteins or other types of molecules. See, Kim *et al.*, 1998 *J. Biochem.* (Tokyo) 124: 1072-1076.

Inhibitors of the interaction between Zmax1 or HBM and interacting proteins may be isolated by standard drug-screening techniques. For example, Zmax1 protein, (or a fragment thereof) or HBM protein (or a fragment thereof) can be immobilized on a solid support such as the base of microtiter well. A second protein or protein fragment, such as ApoE is derivatized to aid in detection, for example with fluorescein. Iodine, or biotin, then added to the Zmax1 or HBM in the presence of candidate compounds that may specifically inhibit this protein-protein domain of Zmax1 or HBM, respectively, and thus avoid problems associated with its transmembrane segment. Drug screens of this type are well known to one skilled in the art of pharmaceutical development.

Because Zmax1 and HBM are involved in bone development, proteins that bind to Zmax1 and HBM are also expected to be involved in bone development. Such binding proteins can be identified by standard methods, such as co-immunoprecipitation, co-fractionation, or the two-hybrid screen (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). For example, to identify Zmax1-interacting proteins or HBM-interacting proteins using the two-hybrid system, the extracellular domain of Zmax1 or HBM is fused to LexA and expressed for the yeast vector pEG202 (the "bait") and expressed in the yeast strain EGY48. The yeast strain is transformed with a "prey" library in the

appropriate vector, which encodes a galactose-inducible transcription-activation sequence fused to candidate interacting proteins. The techniques for initially selecting and subsequently verifying interacting proteins by this method are well known to one skilled in the art of molecular biology (Ausubel *et al.*, 1997).

5 In a preferred embodiment, proteins that interact with HBM, but not Zmax1, are identified using a variation of the above procedure (Xu *et al.*, 1997 *Proc. Natl. Acad. Sci. USA*, 94: 12473-8). This variation of the two-hybrid system uses two baits, and Zmax1 and HBM are each fused to LexA and TetR, respectively. Alternatively, proteins that interact with the HBM but not Zmax1 are also isolated. These procedures are well known to one
10 skilled in the art of molecular biology, and are a simple variation of standard two-hybrid procedures.

As an alternative method of isolating substances interacting with Zmax1 or HBM, a biochemical approach is used. The Zmax1 protein or a fragment thereof, such as the extracellular domain, or the HBM protein or a fragment thereof, such as the extracellular
15 domain, is chemically coupled to Sepharose beads. The Zmax1- or HBM-coupled beads are poured into a column. A biological extract, such as a lipid fraction, serum proteins, proteins in the supernatant of a bone biopsy, or cellular contents from gently lysed osteoblast cells, is added to the column. Non-specifically bound compounds are eluted, the column is washed several times with a low-salt buffer, and then tightly binding compounds may be eluted with
20 a high-salt buffer. These are candidate compounds that bind to Zmax1 or HBM, and can be tested for specific binding by standard tests and control experiments. Sepharose beads used for coupling proteins and the methods for performing the coupling are commercially available (Sigma), and the procedures described here are well known to one skilled in the art of protein biochemistry.

25 As a variation of the above procedure, proteins that are eluted by high salt from the Zmax1- or HBM-sepharose column are then added to an HBM-Zmax1-sepharose column. Proteins that flow through without sticking are proteins that bind to Zmax1 but not to HBM. Alternatively, proteins that bind to the HBM protein and not to the Zmax1 protein can be isolated by reversing the order in which the columns are used.

Isolated compounds may be identified by standard methods such as 2D gel electrophoresis, chromatography, and mass spectroscopy.

24. Method of Use: Transformation-Associated Recombination (TAR) Cloning

5 Essential for the identification of novel allelic variants of Zmax1 (LRP5) and LRP6 is the ability to examine the sequence of both copies of the gene in an individual. To accomplish this, two "hooks," or regions of significant similarity, are identified within the genomic sequence such that they flank the portion of DNA that is to be cloned. Most preferably, the first of these hooks is derived from sequences 5' to the first exon of interest and the second is derived from sequences 3' to the last exon of interest. These two "hooks" are cloned into a bacterial/yeast shuttle vector such as that described by Larionov *et al.*, 1997 *Proc. Natl. Acad. Sci. USA*, 94: 7384-7. Other similar vector systems may also be used. To recover the entire genomic copy of the *Zmax1* gene, the plasmid containing the two "hooks" is linearized with a restriction endonuclease or is produced by another method such as PCR. 10 This linear DNA fragment is introduced into yeast cells along with human genomic DNA. Typically, the yeast *Saccharomyces cerevisiae* is used as a host cell, although Kourina *et al.*, 1998 *Genome Res.* 8: 666-72 have reported using chicken host cells as well. During and after the process of transformation, the endogenous host cell converts the linear plasmid to a circle by a recombination event whereby the region of the human genomic DNA homologous to the "hooks" is inserted into the plasmid. This plasmid can be recovered and analyzed by methods well known to one skilled in the art. Obviously, the specificity for this reaction requires the host cell machinery to recognize sequences similar to the "hooks" present in the linear fragment. However, 100% sequence identity is not required, as shown by Kouprina *et al.*, 1998 *Genomics* 53: 21-8, where the author describes using degenerate repeated sequences common in the human genome to recover fragments of human DNA from a rodent/human 25 hybrid cell line.

 In another example, only one "hook" is required, as described by Larionov *et al.*, 1998 *Proc. Natl. Acad. Sci. USA*, 95: 4469-74. For this type of experiment, termed "radial TAR cloning," the other region of sequence similarity to drive the recombination is derived from a

repeated sequence from the genome. In this way, regions of DNA adjacent to the *Zmax1* gene coding region can be recovered and examined for alterations that may affect function.

25. **Methods of Use: Genomic Screening**

The use of polymorphic genetic markers linked to the *HBM* gene, or to *Zmax1* (LRP5) or to LRP6 is very useful in predicting susceptibility to osteoporosis or other bone diseases. Koller *et al.*, 1998 *Amer. J. Bone Min. Res.* 13: 1903-8 have demonstrated that the use of polymorphic genetic markers is useful for linkage analysis. Similarly, the identification of polymorphic genetic markers within the high bone mass gene will allow the identification of specific allelic variants that are in linkage disequilibrium with other genetic lesions that affect bone development. Using the DNA sequence from the BACs, a dinucleotide CAn repeat was identified and two unique PCR primers that will amplify the genomic DNA containing this repeat were designed, as shown below:

B200E21C16_L: GAGAGGCTATATCCCTGGGC (SEQ ID NO: 38)

B200E21C16_R: ACAGCACGTGTTTAAAGGGG (SEQ ID NO: 39)

and used in the genetic mapping study.

This method has been used successfully by others skilled in the art (e.g., Sheffield *et al.*, 1995 *Genet.*, 4:1837-44; LeBlanc-Straceski *et al.*, 1994 *Genomics*, 19: 341-9; Chen *et al.*, 1995 *Genomics*, 25:1-8). Use of these reagents with populations or individuals will predict their risk for osteoporosis. Similarly, single nucleotide polymorphisms (SNPs), such as those shown in Table 4 above, can be used as well to predict risk for developing bone diseases or resistance to osteoporosis in the case of the *HBM* gene.

26. **Methods of Use: Modulators of Tissue Calcification**

The calcification of tissues in the human body is well documented. Towler *et al.* (1998 *J. Biol. Chem.* 273: 30427-34) demonstrated that several proteins known to regulate calcification of the developing skull in a model system are expressed in calcified aorta. The expression of *Mx2*, a gene transcribed in osteoprogenitor cells, in calcified vascular tissue indicates that genes which are important in bone development are involved in calcification of

other tissues.

Treatment with HBM protein, HBM-like proteins and polypeptides, agonists of LRP5 and LRP6 and HBM or antagonists of Dkk-1 are likely to ameliorate calcification (such as the vasculature, dentin and bone of the skull viscera) due to its demonstrated effect on bone mineral density. In experimental systems where tissue calcification is demonstrated, the over-expression or repression of *Zmax1*(LRP5) activity permits the identification of molecules that are directly regulated by the *Zmax1* gene. These genes are potential targets for therapeutics aimed at modulating tissue calcification. For example, an animal, such as the LDLR ^{-/-}, mouse is fed a high fat diet and is observed to demonstrate expression of markers of tissue calcification, including *Zmax1*. These animals are then treated with antibodies to *Zmax1* or HBM protein, antisense oligonucleotides directed against *Zmax1* or HBM cDNA, or with compounds known to bind the *Zmax1* or HBM protein or its binding partner or ligand. RNA or proteins are extracted from the vascular tissue and the relative expression levels of the genes expressed in the tissue are determined by methods well known in the art. Genes that are regulated in the tissue are potential therapeutic targets for pharmaceutical development as modulators of tissue calcification.

The nucleic acids, proteins, peptides, amino acids, small molecules or other pharmaceutically useful compounds of the present invention that are to be given to an individual may be administered in the form of a composition with a pharmaceutically acceptable carrier, excipient or diluent, which are well known in the art. The individual may be a mammal or a bird, preferably a human, a rat, a mouse or bird. Such compositions may be administered to an individual in a pharmaceutically effective amount. The amount administered will vary depending on the condition being treated and the patient being treated. The compositions may be administered alone or in combination with other treatments.

27. Methods to Identify Agents that Modulate the Expression of a Nucleic Acid Encoding the Dkk and/or LRP5 Proteins and/or Dkk interacting proteins

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding Dkk, which is part of the Wnt

pathway and that interacts with LRP5, LRP6 and to much lesser extent to HBM and its variants. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of Dkk, if it is capable of up- or down-regulating expression of the nucleic acid in a cell (*e.g.*, mRNA).

In one assay format, cell lines that contain reporter gene fusions between the nucleic acid encoding Dkk (or proteins which modulate the activity of Dkk) and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available, including but not limited to the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.*, 1990 *Anal. Biochem.* 188: 245-54). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding Dkk or other protein which modulates Dkk activity. Such assays can similarly be used to determine whether LRP5 and even LRP6 activity is modulated by regulating Dkk activity. This can also be performed with the HBM variants.

Additional assay formats may be used to monitor the ability of the agent(s) to modulate the expression of a nucleic acid encoding Dkk, alone or Dkk and LRP5, and/or Dkk interacting proteins such as those identified in Figure 31. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.* (1989); Ausubel *et al.*, Current Protocols in Molecular Biology (Greene Publishing Co., NY, 1995); Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982); and Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology (Frederick M. Ausubel *et al.*, April 1999).

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is

preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available. See for example, Sambrook *et al.* (1989) or Ausubel *et al.* (Current Protocols in Molecular Biology, Greene Publishing Co., NY, 1995).

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.* (1989) and Ausubel *et al.* (1995), as suitable for each probe.

Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the nucleic acid sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass or silica wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up- or down-regulate the expression of a nucleic acid encoding Dkk, a Dkk interacting protein, and/or LRP5 can be identified.

Microarray technology and transcriptional profiling are examples of methods which can be used to analyze the impact of putative Dkk or Dkk interacting protein modulating compounds. For transcriptional profiling, mRNA from cells exposed *in vivo* to a potential

Dkk modulating agent, such as the Dkk interacting proteins identified in the present invention (e.g., those identified in Fig. 31), agents which modulate Dkk interacting proteins, and mRNA from the same type of cells that were not exposed to the agent could be reverse transcribed and hybridized to a chip containing DNA from numerous genes, to thereby
5 compare the expression of genes in cells treated and not treated with the agent. If, for example a putative Dkk modulating agent down-regulates the expression of Dkk in the cells, then use of the agent may be undesirable in certain patient populations. For additional methods of transcriptional profiling and the use of microarrays, refer to, for example, U.S. Patent No. 6,124,120 issued to Lizardi (2000).

10 Additional methods for screening the impact of Dkk and Dkk interacting protein modulating compounds or the impact of Dkk or Dkk interacting proteins on modulation of LRP5, LRP6, HBM, HBM variants or the Wnt pathway include the use of TaqMan® PCR, conventional reverse transcriptase PCR (RT-PCR), changes in downstream surrogate markers (i.e., Wnt responsive genes), and anti-Dkk Western blots for protein detection. Other
15 methods would be readily apparent to the artisan of ordinary skill.

28. Methods to Identify Agents that Modulate at Least One Activity of Dkk, a Dkk Interacting Protein, or LRP5/LRP6/HBM/HBM-like

Another embodiment of the present invention provides methods for identifying agents
20 that modulate at least one activity of Dkk, Dkk interacting proteins, and/or LRP5/LRP6/HBM/HBM-like proteins or preferably which specifically modulate an activity of a Dkk/Dkk interacting protein complex or an LRP5 (or LRP6/HBM/HBM-like)/Dkk complex, or a biologically active fragment of Dkk (e.g., comprising the domain which binds LRP5/LRP6/HBM/HBM-like) or a Dkk interacting protein complex. Such methods or
25 assays may utilize any means of monitoring or detecting the desired activity as would be known in the art (See, e.g., Wu *et al.*, 2000 *Curr. Biol.* 10:1611-4; Fedi *et al.*, 199 *J. Biol. Chem.* 274:19465-72; Grotewold *et al.*, 1999 *Mech. Dev.* 89:151-3; Shibata *et al.*, 2000 *Mech. Dev.* 96: 243-6 ; Wang *et al.*, 2000 *Oncogene* 19: 1843-8; and Glinka *et al.*, 1998 *Nature* 391: 357-62). Potential agents which modulate Dkk include, for example, p53, the

tumor suppressor protein, which can induce Dkk-1. Damage to DNA has also been observed to up-regulate Dkk-1 expression via a stabilization and activation of p53 (Wang *et al.*, 2000 *Oncogene* 19: 1843-48); and, Shou *et al.*, 2002 *Oncogene* 21: 878-89). Additionally, Fedi *et al.* (1999) purportedly showed that Dkk-1 can block the Wnt2-induced oncogenic transformation of NIH-3T3 cells. Furthermore, it has been suggested that Dkk expression may be modulated by BMP signaling in the developing skeleton (Mukhopadhyay *et al.*, 2001 *Dev. Cell.* 1: 423-34; and Grotewold *et al.*, 2002 *EMBO J.* 21: 966-75). Grotewald *et al.* additionally describe altered Dkk expression levels in response to stress signals including UV irradiation and other genotoxic stimuli. They propose that Dkk expression is pro-apoptotic. In the HBMMTIC animals described herein, a reduced osteoblast apoptosis effect was observed. Thus, HBM and HBM like variants may control/alter Dkk's role in programmed cell death. Other agents which potentially modulate Dkk activity include the Dkk interacting proteins identified in Fig. 31.

In one embodiment, the relative amounts of Dkk or a Dkk interacting protein of a cell population that has been exposed to the agent to be tested is compared to an un-exposed control cell population. Antibodies can be used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe, as would be known in the art. *See, e.g.*, Ed Harlow and David Lane, Antibodies: A Laboratory Manual (Cold Spring Harbor, NY, 1988) and Ed Harlow and David Lane, Using Antibodies: A Laboratory Manual (Cold Spring Harbor, NY 1998).

For example, N- and C- terminal fragments of Dkk can be expressed in bacteria and used to search for proteins which bind to these fragments. Fusion proteins, such as His-tag or GST fusion to the N- or C-terminal regions of Dkk (or to biologically active domains of Dkk-1) or a whole Dkk protein can be prepared. These fusion proteins can be coupled to, for example, Talon or Glutathione-Sepharose beads and then probed with cell lysates to identify molecules which bind to Dkk. Prior to lysis, the cells may be treated with purified Wnt

proteins, RNA, or drugs which may modulate Wnt signaling or proteins that interact with downstream elements of the Wnt pathway. Lysate proteins binding to the fusion proteins can be resolved by SDS-PAGE, isolated and identified by, for example protein sequencing or mass spectroscopy, as is known in the art. See, e.g., Protein Purification Applications: A Practical Approach (Simon Roe, ed., 2nd ed. Oxford Univ. Press, 2001) and "Guide to Protein Purification" in *Meth. Enzymology* vol. 182 (Academic Press, 1997).

The activity of Dkk, a Dkk interacting protein, or a complex of Dkk with LRP5/LRP6/HBM/HBM-like may be affected by compounds which modulate the interaction between Dkk and a Dkk interacting protein (such as those shown in Fig. 31) and/or Dkk and LRP5/LRP6/HBM/HBM-like. The present invention provides methods and research tools for the discovery and characterization of these compounds. The interaction between Dkk and a Dkk interacting protein and/or Dkk and LRP5/6/HBM/HBM-like may be monitored *in vivo* and *in vitro*. Compounds which modulate the stability of a Dkk - LRP5/LRP6/HBM/HBM-like complex are potential therapeutic compounds. Example *in vitro* methods include:

Binding LRP5/6/HBM/HBM-like, Dkk, or a Dkk interacting protein to a sensor chip designed for an instrument such are made by Biacore (Uppsala, Sweden) for the performance of an plasmon resonance spectroscopy observation. In this method, the chip with one of Dkk, a Dkk interacting protein, or LRP5/6 is first exposed to the other under conditions which permit them to form the complex. A test compound is then introduced and the output signal of the instrument provides an indication of any effect exerted by the test compound. By this method, compounds may be rapidly screened. Another, *in vitro*, method is exemplified by the SAR-by-NMR methods (Shuker *et al.*, *Science*. 274:1531-4 (1996)). Briefly, a Dkk-1 binding domain and/or LRP 5 or 6 LBD are expressed and purified as ¹⁵N labeled protein by expression in labeled media. The labeled protein(s) are allowed to form the complex in solution in an NMR sample tube. The heteronuclear correlation spectrum in the presence and absence of a test compound provides data at the level of individual residues with regard to interactions with the test compound and changes at the protein-protein interface of the complex. One of skill in the art knows of many other protocols, e.g. affinity capillary electrophoresis (Okun *et al.*, 2001 *J. Biol. Chem.* 276: 1057-62; Vergnon *et al.*, 1999

Methods, 19: 270-7), fluorescence spectroscopy, electron paramagnetic resonance, etc. which can monitor the modulation of a complex and/or measure binding affinities for complex formation.

In vitro protocols for monitoring the modulation of a Dkk/LRP5/LRP6/HBM/HBM-like complex include the yeast two hybrid protocol. The yeast two hybrid method may be used to monitor the modulation of a complex *in vivo* by monitoring the expression of genes activated by the formation of a complex of fusion proteins of Dkk and LRP ligand binding domains. Nucleic acids according to the invention which encode the interacting Dkk and LRP LBD domains are incorporated into bait and prey plasmids. The Y2H protocol is performed in the presence of one or more test compounds. The modulation of the complex is observed by a change in expression of the complex activated gene. It will be appreciated by one skilled in the art that test compounds can be added to the assay directly or, in the case of proteins, can be coexpressed in the yeast with the bait and prey compounds. Similarly, fusion proteins of Dkk and Dkk interacting proteins can also be used in a Y2H screen to identify other proteins which modulate the Dkk/Dkk interacting protein complex.

Assay protocols such as these may be used in methods to screen for compounds, drugs, treatments which modulate the Dkk/Dkk interacting protein and/or Dkk/LRP5/6 complex, whether such modulation occurs by competitive binding, or by altering the structure of either LRP 5/6 or Dkk at the binding site, or by stabilizing or destabilizing the protein-protein interface. It may be anticipated that peptide aptamers may competitively bind, although induction of an altered binding site structure by steric effects is also possible.

28.1 Antibodies and Antibody Fragments

Polyclonal and monoclonal antibodies and fragments of these antibodies which bind to Dkk or LRP5/LRP6/HBM/HBM-like can be prepared as would be known in the art. For example, suitable host animals can be immunized using appropriate immunization protocols and the peptides, polypeptides or proteins of the invention. Peptides for use in immunization are typically about 8-40 residues long. If necessary or desired, the polypeptide immunogens can be conjugated to suitable carriers. Methods for preparing immunogenic conjugates with

carriers such as bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), or other carrier proteins are well known in the art (*See, Harlow et al.*, 1988). In some circumstances, direct conjugation using, for example, carbodiimide reagents, may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, IL, may be desirable to provide accessibility to the polypeptide or hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

Anti-peptide antibodies can be generated using synthetic peptides, for example, the peptides derived from the sequence of any Dkk, including Dkk-1, or LRP5/LRP6/HBM/HBM-like. Synthetic peptides can be as small as 2-3 amino acids in length, but are preferably at least 3, 5, 10, or 15 or more amino acid residues long. Such peptides can be determined using programs such as DNASTar. The peptides are coupled to KLH using standard methods and can be immunized into animals such as rabbits. Polyclonal anti-Dkk or anti-LRP5/LRP6/HBM/HBM-like peptide antibodies can then be purified, for example using Actigel beads containing the covalently bound peptide.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known (*See, e.g., Harlow et al.*, 1988 and 1998). The immortalized cell lines secreting the desired antibodies can be screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies which contain the

immunologically significant portion can be used as agonists or antagonists of Dkk activity. Use of immunologically reactive fragments, such as the Fab, scFV, Fab', of F(ab')₂ fragments are often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

5 The antibodies or fragments may also be produced, using current technology, by recombinant means. Regions that bind specifically to the desired regions of Dkk or LRP5/LRP6/HBM/HBM-like can also be produced in the context of chimeras with multiple species origin. Antibody reagents so created are contemplated for use diagnostically or as stimulants or inhibitors of Dkk activity.

10 In one embodiment, antibodies against Dkk, bind Dkk with high affinity, i.e., ranging from 10⁻⁵ to 10⁻⁹ M. Preferably, the anti-Dkk antibody will comprise a chimeric, primate, Primatized®, human or humanized antibody. Also, the invention embraces the use of antibody fragments, e.g., Fab's, Fv's, Fab's, F(ab)₂, and aggregates thereof.

15 Another embodiment contemplates chimeric antibodies which recognize Dkk or LRP5/LRP6/HBM/HBM-like. A chimeric antibody is intended to refer to an antibody with non-human variable regions and human constant regions, most typically rodent variable regions and human constant regions.

20 A "primatized® antibody" refers to an antibody with primate variable regions, e.g., CDR's, and human constant regions. Preferably, such primate variable regions are derived from an Old World monkey.

25 A "humanized antibody" refers to an antibody with substantially human framework and constant regions, and non-human complementarity-determining regions (CDRs). "Substantially" refers to the fact that humanized antibodies typically retain at least several donor framework residues (i.e., of non-human parent antibody from which CDRs are derived).

 Methods for producing chimeric, primate, primatized®, humanized and human antibodies are well known in the art. See, e.g., U.S. Patent 5,530,101, issued to Queen *et al.*; U.S. Patent 5,225,539, issued to Winter *et al.*; U.S. Patents 4,816,397 and 4,816,567, issued to Boss *et al.* and Cabilly *et al.* respectively, all of which are incorporated by reference in

their entirety.

The selection of human constant regions may be significant to the therapeutic efficacy of the subject anti-Dkk or LRP5/LRP6/HBM/HBM-like antibody. In a preferred embodiment, the subject anti-Dkk or LRP5/LRP6/HBM/HBM-like antibody will comprise human, gamma 1, or gamma 3 constant regions and, more preferably, human gamma 1 constant regions.

Methods for making human antibodies are also known and include, by way of example, production in SCID mice, and *in vitro* immunization.

The subject anti-Dkk or LRP5/LRP6/HBM/HBM-like antibodies can be administered by various routes of administration, typically parenteral. This is intended to include intravenous, intramuscular, subcutaneous, rectal, vaginal, and administration with intravenous infusion being preferred.

The anti-Dkk or LRP5/LRP6/HBM/HBM-like antibody will be formulated for therapeutic usage by standard methods, *e.g.*, by addition of pharmaceutically acceptable buffers, *e.g.*, sterile saline, sterile buffered water, propylene glycol, and combinations thereof.

Effective dosages will depend on the specific antibody, condition of the patient, age, weight, or any other treatments, among other factors. Typically effective dosages will range from about 0.001 to about 30 mg/kg body weight, more preferably from about 0.01 to 25 mg/kg body weight, and most preferably from about 0.1 to about 20 mg/kg body weight.

Such administration may be effected by various protocols, *e.g.*, weekly, bi-weekly, or monthly, depending on the dosage administered and patient response. Also, it may be desirable to combine such administration with other treatments.

Antibodies to Dkk-1 interacting proteins, such as those identified in Fig. 31, are also contemplated according to the present invention, and can be used similarly to the Dkk-1 antibodies mentioned in the above methodology.

The antibodies of the present invention can be utilized in experimental screening, as diagnostic reagents, and in therapeutic compositions.

28.2 Chemical Libraries

Agents that are assayed by these methods can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of Dkk-1 alone, Dkk-1 interacting proteins alone, or with their associated substrates, binding partners, etc. An example of randomly selected agents is the use of a chemical library or a peptide combinatorial library, or a growth broth of an organism.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

28.3 Peptide Synthesis

The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production of polypeptides using solid phase peptide synthesis is necessitated if non-nucleic acid-encoded amino acids are to be included.

29. Uses for Agents that Modulate at Least One Activity of Dkk, a Dkk Interacting Protein, a Dkk/Dkk Interacting Protein Complex, or a Dkk/LRP5 or Dkk/LRP6 Complex

The proteins and nucleic acids of the invention, such as the proteins or polypeptides containing an amino acid sequence of LRP5, Dkk, and Dkk interacting proteins are involved in bone mass modulation and lipid modulation of other Wnt pathway mediated activity.

Agents that modulate (*i.e.*, up and down-regulate) the expression of Dkk or Dkk interacting proteins, or agents, such as agonists and antagonists respectively, of at least one activity of Dkk or a Dkk interacting protein may be used to modulate biological and pathologic processes associated with the function and activity of Dkk or a Dkk interacting protein.

As used herein, a subject can be preferably any mammal, so long as the mammal is in

need of modulation of a pathological or biological process modulated by a protein of the invention. The term “mammal” means an individual belonging to the class *Mammalia*. The invention is particularly useful in the treatment of human subjects.

As used herein, a biological or pathological process modulated by Dkk or a Dkk
5 interacting protein may include binding of Dkk to a Dkk interacting protein, Dkk to LRP5 or LRP6 or release therefrom, inhibiting or activating Dkk or a Dkk interacting protein mRNA synthesis or inhibiting Dkk or Dkk interacting protein modulated inhibition of LRP5 or LRP6 mediated Wnt signaling. Further bone-related markers may be observed such as alkaline phosphatase activity, osteocalcin production, or mineralization.

10 Pathological processes refer to a category of biological processes which produce a deleterious effect. For example, expression or up-regulation of expression of LRP5 or LRP6 and/or Dkk and/or a Dkk interacting protein may be associated with certain diseases or pathological conditions. As used herein, an agent is said to modulate a pathological process when the agent statistically significantly ($p < 0.05$) alters the process from its base level in
15 the subject. For example, the agent may reduce the degree or severity of the process mediated by that protein in the subject to which the agent was administered. For instance, a disease or pathological condition may be prevented, or disease progression modulated by the administration of agents which reduce or modulate in some way the expression or at least one activity of a protein of the invention.

20 As LRP5/6 and Dkk are involved both directly and indirectly in bone mass modulation, one embodiment of this invention is to use Dkk or Dkk interacting protein expression as a method of diagnosing a bone condition or disease. Certain markers are associated with specific Wnt signaling conditions (e.g., *TCF/LEF* activation). Diagnostic tests for bone conditions may include the steps of testing a sample or an extract thereof for
25 the presence of Dkk or Dkk interacting protein nucleic acids (i.e., DNA or RNA), oligomers or fragments thereof or protein products of *TCF/LEF* regulated expression. For example, standard *in situ* hybridization or other imaging techniques can be utilized to observe products of Wnt signaling. Other diagnostic techniques, as described herein, would also be useful as would be apparent to the skilled artisan (e.g., a serum marker).

This invention also relates to methods of modulating bone development or bone loss conditions. Inhibition of bone loss may be achieved by inhibiting or modulating changes in the LRP5/6 mediated Wnt signaling pathway. For example, absence of LRP5 activity may be associated with low bone mass. Increased activity LRP5 may be associated with high bone mass. Therefore, modulation of LRP5 activity will in turn modulate bone development. Modulation of the Dkk/LRP5/6 or Dkk/Dkk interacting protein complex via agonists and antagonists is one embodiment of a method to regulate bone development. Such modulation of bone development can result from inhibition of the activity of, for example, a Dkk/LRP(5/6) protein complex, a Dkk/Dkk interacting protein complex, upregulated transcription of the *LRP5* gene or inhibited transcription or translation of Dkk or Dkk interacting protein mRNA.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular pathological process. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention can be administered via parenteral, subcutaneous (sc), intravenous (iv), intramuscular (im), intraperitoneal (ip), transdermal or buccal routes. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

The present invention further provides compositions containing one or more agents which modulate expression or at least one activity of a protein of the invention. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typical dosages of the active agent which mediate Dkk or Dkk interacting protein activity comprise from about 0.0001 to about 50 mg/kg body weight. The preferred dosages comprise from about 0.001 to about 50 mg/kg body weight. The most preferred dosages comprise from about 0.1 to about 1 mg/kg body weight. In an average human of 70 kg, the range would be from about 7 μ g to about 3.5 g, with a preferred range of about 0.5 mg to about 5 mg.

In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, (e.g., ethyl oleate or triglycerides). Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol and/or dextran. Optionally, the suspension may also contain stabilizers. Liposomes and other non-viral vectors can also be used to encapsulate the agent for delivery into the cell.

The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral, or topical (top) administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

Potentially, any compound which binds Dkk or a Dkk interacting protein or modulates the Dkk/LRP5 or Dkk/LRP6 or Dkk/Dkk interacting protein complex may be a therapeutic compound. In one embodiment of the invention, a peptide or nucleic acid aptamer according to the invention is used in a therapeutic composition. Such compositions may comprise an aptamer, or a LRP5 or LRP6 fragment unmodified or modified. In another embodiment, the therapeutic compound comprises a Dkk-1 interacting protein, or biologically active fragment thereof.

Nucleic acid aptamers have been used in compositions for example by chemical bonding to a carrier molecule such as polyethylene glycol (PEG) which may facilitate uptake

or stabilize the aptamer. A di-alkylglycerol moiety attached to an RNA will embed the aptamer in liposomes, thus stabilizing the compound. Incorporating chemical substitutions (i.e. changing the 2'OH group of ribose to a 2'NH in RNA confers ribonuclease resistance) and capping, etc. can prevent breakdown. Several such techniques are discussed for RNA aptamers in Brody and Gold (*Rev. Mol. Biol.* 74:3-13, 2000).

Peptide aptamers may be used in therapeutic applications by the introduction of an expression vector directing aptamer expression into the affected tissue such as for example by retroviral delivery, by encapsulating the DNA in a delivery complex or simple by naked DNA injection. Or, the aptamer itself or a synthetic analog may be used directly as a drug. Encapsulation in polymers and lipids may assist in delivery. The use of peptide aptamers as therapeutic and diagnostic agents is reviewed by Hoppe-Syler and Butz (*J. Mol. Med.* 78:426-430 (2000)).

In another aspect of the invention. The structure of a constrained peptide aptamer of the invention may be determined such as by NMR or X-ray crystallography. (Cavanagh *et al.*, Protein NMR Spectroscopy: Principles and Practice, Academic Press, 1996; Drenth, Principles of Protein X-Ray Crystallography, Springer Verlag, 1999) Preferably the structure is determined in complex with the target protein. A small molecule analog is then designed according to the positions of functional elements of the 3D structure of the aptamer. (Guidebook on Molecular Modeling in Drug Design, Cohen, Ed., Academic Press, 1996; Molecular Modeling and Drug Design (Topics in Molecular and Structural Biology), Vinter and Gardner Eds., CRC Press, 1994) Thus the present invention provides a method for the design of effective and specific drugs which modulate the activity of Dkk, Dkk interacting proteins, Dkk/Dkk interacting protein complex and the Dkk/LRP complex. Small molecule mimetics of the peptide aptamers of the present invention are encompassed within the scope of the invention.

In practicing the methods of this invention, the compounds of this invention may be used alone or in combination, or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be co-administered along with other compounds typically prescribed for these conditions according to generally

accepted medical practice. For example, the compounds of this invention can be administered in combination with other therapeutic agents for the treatment of bone loss. Bone loss mediating agents include bone resorption inhibitors such as bisphosphonates (*e.g.*, alendronic acid, clodronic acid, etidronic acid, pamidronic acid, risedronic acid and tiludronic acid), vitamin D and vitamin D analogs, cathepsin K inhibitors, hormonal agents (*e.g.*, calcitonin and estrogen), and selective estrogen receptor modulators or SERMs (*e.g.*, raloxifene). And bone forming agents such as parathyroid hormone (PTH) and bone morphogenetic proteins (BMP).

Additionally contemplated are combinations of agents which regulate Dkk-1 and agents which regulate lipid levels such as HMG-CoA reductase inhibitors (*i.e.*, statins such as Mevacor®, Lipitor® and other inhibitors such as Baycol®, Lescol®, Pravachol® and Zocor®), bile acid sequestrants (*e.g.*, Colestid® and Welchol®), fibric acid derivatives (Atromid-S®, Lopid®, Tricor®), and nicotinic acid.

The compounds of this invention can be utilized *in vivo*, ordinarily in vertebrates and preferably in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or *in vitro*.

30. Peptide and Nucleotide Aptamers and Peptide Aptamer Mimetics

Another embodiment contemplates the use of peptide and nucleotide aptamer technology to screen for agents which interact with Dkk, which block Dkk from interacting with LRP5, LRP6, HBM, or HBM-like molecules, or which block any other Dkk ligand interaction, or which interact with Dkk interacting proteins, such as those shown in Figure 5. Peptide aptamers are molecules in which a variable peptide domain is displayed from a scaffold protein. Thioredoxin A (trxA) is commonly used for a scaffold. The peptide insert destroys the catalytic site of trxA. It is recognized that numerous proteins may also be used as scaffolding proteins to constrain and/or present a peptide aptamer. Other scaffold proteins that could display a constrained peptide aptamer could include staphylococcal nuclease, the protease inhibitor eglin C, the *Streptomyces tendae* alpha-amylase inhibitor Tendamistat, Sp1, green fluorescent protein (GFP) (reviewed in Hoppe-Seyler *et al.*, 2001 *J. Steroid*

Biochem Mol. Biol. 78: 105-11), and the S1 nucleas from *Staphyloccus* or M13 for phage display.. Any molecule to which the aptamer could be anchored and presented in its bioactive conformation would be suitable.

Aptamers can then specifically bind to a given target protein *in vitro* and *in vivo* and have the potential to selectively block the function of their target protein. Peptide aptamers are selected from randomized expression libraries on the basis of their *in vivo* binding capacity to the desired target protein. Briefly, a target protein (e.g., Dkk, a Dkk interacting protein, or LRP5/6) is linked to a heterologous DNA binding domain (BD) and expressed as bait in a yeast test strain. Concomitantly, a library coding for different peptides (e.g., 16-mers) of randomized sequence inserted in a scaffold protein sequence, which are linked to a heterologous transcriptional activation domain (AD) is expressed as prey. If a peptide binds to a target protein, a functional transcription factor is reconstituted, in which the BD and AD are bridged together by interacting proteins. This transcription factor is then able to activate the promoter of a marker gene which can be monitored by colorimetric enzymatic assays or by growth selection. Additional variation, methods of preparing and screening methodologies are described in, for example, Hoppe-Seyler *et al.*, 2000 *J. Mol. Med.* 78: 426-430.

Nucleotide aptamers are described for example in Brody *et al.*, 2000 *Trends Mol. Biotechnol.* 74: 5-13. Additional methods of making and using nucleotide aptamers include SELEX, *i.e.*, Systematic Evolution of Ligands by Exponential Enrichment. SELEX is a process of isolating oligonucleotide ligands of a chosen target molecule (see Tuerk and Gold, *Science* 249:505-510 (1990); U.S. Pat. Nos. 5,475,096, 5,595,877, and 5,660,985). SELEX, as described in Tuerk and Gold, involves admixing the target molecule with a pool of oligonucleotides (e.g., RNA) of diverse sequences; retaining complexes formed between the target and oligonucleotides; recovering the oligonucleotides bound to the target; reverse-transcribing the RNA into DNA; amplifying the DNA with polymerase chain reactions (PCR); transcribing the amplified DNA into RNA; and repeating the cycle with ever increasing binding stringency. Three enzymatic reactions are required for each cycle. It usually takes 12-15 cycles to isolate aptamers of high affinity and specificity to the target.

An aptamer is an oligonucleotide that is capable of binding to an intended target substance but not other molecules under the same conditions.

In another reference, Bock *et al.*, 1990 *Nature* 355: 564-6, describe a different process from the SELEX method of Tuerk and Gold in that only one enzymatic reaction is required for each cycle (*i.e.*, PCR) because the nucleic acid library in Bock's method is comprised of DNA instead of RNA. The identification and isolation of aptamers of high specificity and affinity with the method of Bock *et al.* still requires repeated cycles in a chromatographic column.

Other nucleotide aptamer methods include those described by Conrad *et al.*, 1996 *Meth. Enzymol.* 267: 336-367. Conrad *et al.* describe a variety of methods for isolating aptamers, all of which employ repeated cycles to enrich target-bound ligands and require a large amount of purified target molecules. More recently described methods of making and using nucleotide aptamers include, but are not limited to those described in U.S. Patent Nos. 6,180,348; 6,051,388; 5,840,867; 5,780,610, 5,756,291 and 5,582,981.

Potentially, any compound which binds Dkk or a Dkk interacting protein or modulates the Dkk/Dkk interacting protein or Dkk/LRP5, Dkk/LRP6, Dkk/HBM, or Dkk/HBM-like complex may be a therapeutic compound. In one embodiment of the invention, a peptide or nucleic acid aptamer according to the invention is used in a therapeutic composition. Such compositions may comprise an aptamer, or a LRP5 or LRP6 fragment unmodified or modified.

Nucleic acid aptamers have been used in compositions for example by chemical bonding to a carrier molecule such as polyethylene glycol (PEG) which may facilitate uptake or stabilize the aptamer. A di-alkylglycerol moiety attached to an RNA will embed the aptamer in liposomes, thus stabilizing the compound. Incorporating chemical substitutions (*i.e.*, changing the 2'-OH group of ribose to a 2'-NH in RNA confers ribonuclease resistance) and capping, etc. can prevent breakdown. Several such techniques are discussed for RNA aptamers in Brody *et al.*, 2000 *Rev. Mol. Biol.* 74: 3-13.

Peptide aptamers may be used in therapeutic applications by the introduction of an expression vector directing aptamer expression into the affected tissue such as for example by

retroviral delivery, by encapsulating the DNA in a delivery complex or simple by naked DNA injection. Or, the aptamer itself or a synthetic analog may be used directly as a drug. Encapsulation in polymers and lipids may assist in delivery. The use of peptide aptamers as therapeutic and diagnostic agents is reviewed by Hoppe-Syler *et al.*, 2000 *J. Mol. Med.* 78: 426-430.

In another aspect of the invention, the structure of a constrained peptide aptamer of the invention may be determined such as by NMR or X-ray crystallography. (Cavanagh *et al.*, Protein NMR Spectroscopy : Principles and Practice, Academic Press, 1996; Drenth, Principles of Protein X-Ray Crystallography, Springer Verlag, 1999) Preferably the structure is determined in complex with the target protein. A small molecule analog is then designed according to the positions of functional elements of the 3D structure of the aptamer. (Guidebook on Molecular Modeling in Drug Design, Cohen, Ed., Academic Press, 1996; Molecular Modeling and Drug Design (Topics in Molecular and Structural Biology), Vinter and Gardner Eds., CRC Press, 1994) Thus, a method is provided for the design of effective and specific drugs which modulate the activity of Dkk, Dkk interacting proteins, Dkk/Dkk interacting protein complex, and the Dkk/LRP complex. Small molecule mimics of the peptide aptamers of the present invention are also encompassed within the scope of the invention.

EXAMPLES

The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well known in the art or the techniques specifically described below were utilized.

Example 1

The propositus was referred by her physicians to the Creighton Osteoporosis Center for evaluation of what appeared to be unusually dense bones. She was 18 years old and came to medical attention two years previous because of back pain, which was precipitated by an

auto accident in which the car in which she was riding as a passenger was struck from behind. Her only injury was soft tissue injury to her lower back that was manifested by pain and muscle tenderness. There was no evidence of fracture or subluxation on radiographs. The pain lasted for two years, although she was able to attend school full time. By the time
5 she was seen in the Center, the pain was nearly resolved and she was back to her usual activities as a high school student. Physical exam revealed a normal healthy young woman standing 66 inches and weighing 128 pounds. Radiographs of the entire skeleton revealed dense looking bones with thick cortices. All bones of the skeleton were involved. Most importantly, the shapes of all the bones were entirely normal. The spinal BMC was 94.48
10 grams in L1-4, and the spinal BMD was 1.667 gm/cm² in L1-4. BMD was 5.62 standard deviations (SD) above peak skeletal mass for women. These were measured by DXA using a Hologic 2000~. Her mother was then scanned and a lumbar spinal BMC of 58.05 grams and BMD of 1.500 gm/cm² were found. Her mother's values place her 4.12 SD above peak mass and 4.98 SD above her peers. Her mother was 51 years old, stood 65 inches and weighed
15 140 pounds. Her mother was in excellent health with no history of musculoskeletal or other symptoms. Her father's lumbar BMC was 75.33 grams and his BMD was 1.118 gm/cm². These values place him 0.25 SD above peak bone mass for males. He was in good health, stood 72 inches tall, and weighed 187 pounds.

These clinical data suggested that the propositus inherited a trait from her mother,
20 which resulted in very high bone mass, but an otherwise normal skeleton, and attention was focused on the maternal kindred. In U.S. Patent No. 5,691,153, twenty- two of these members had measurement of bone mass by DXA. In one case, the maternal grandfather of the propositus, was deceased, however, medical records, antemortem skeletal radiographs and a gall bladder specimen embedded in paraffin for DNA genotyping were obtained. His
25 radiographs showed obvious extreme density of all of the bones available for examination including the femur and the spine, and he was included among the affected members. In this invention, the pedigree has been expanded to include 37 informative individuals. These additions are a significant improvement over the original kinship (Johnson *et al.*, *Am. J. Hum. Genet.*, 60:1326-1332 (1997)) because, among the fourteen individuals added since the

original study, two individuals harbor key crossovers. X-linkage is ruled out by the presence of male-to-male transmission from individual 12 to 14 and 15.

Example 2

5 The present invention describes DNA sequences derived from two BAC clones from the *HBM* gene region, as evident in Table 7 below, which is an assembly of these clones. Clone b200e21-h (ATCC No. 980812; SEQ ID NOS: 10-11) was deposited at the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 U.S.A., on December 30, 1997. Clone b527d12-h (ATCC No. 980720; SEQ ID NOS: 5-9)
10 was deposited at the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 U.S.A., on October 2, 1998. These sequences are unique reagents that can be used by one skilled in the art to identify DNA probes for the *Zmax1* gene, PCR primers to amplify the gene, nucleotide polymorphisms in the *Zmax1* gene, or regulatory elements of the *Zmax1* gene.

Example 3

Yeast-2 Hybrid screen for peptide aptamer sequences to Dkk-1

Peptide aptamer library construction. A peptide aptamer library, Tpep, was constructed, which provides a means to identify chimeric proteins that bind to a protein target
20 (or bait) of interest using classic yeast two hybrid (Y2H) assays. The Tpep library is a combinatorial aptamer library composed of constrained random peptides, expressed within the context of the disulfide loop of *E. coli* thioredoxin (trxA), and as C-termini fusion to the *S. cerevisiae* Gal4 activation domain. The Tpep library was generated using a restriction enzyme modified recombinant Y2H prey vector, pPC86 (Gibco), which contains the trxA
25 scaffold protein.

Generation of aptamer-encoding sequences. Aptamer-encoding sequences were produced as follows. DNA encoding random stretches of approximately sixteen amino acids surrounded by appropriate restriction sites were generated by semi-random oligonucleotide

synthesis. The synthetic oligonucleotides were PCR-amplified, restriction digested, and cloned into the permissive sites within the *trxA* scaffold protein. The cloning strategy was to insert the random oligonucleotide sequence in-frame with the scaffold protein coding sequence, resulting in expression of a scaffold protein-aptamer chimera. The scaffold protein is itself in-frame with the activation domain of Gal4, within the pPC86 vector that is appropriate for the aptamer to be expressed and functional in a regular Y2H assay. Additional methods of preparing aptamers would be apparent to the skilled artisan.

Generation of a permissive recombinant pPC86 vector containing the *trxA* coding sequence. First the *Rsr*II restriction site located within the Gal4 activation domain of pPC86 (Gibco) was eliminated by site-directed mutagenesis (Quickchange™ kit, Stratagene). The amino acid sequence of the Gal4 activation domain was unchanged by this modification. The strength of different control interactions was verified to be unchanged by the modification.

Second, the *E. coli* *trxA* coding sequence was cloned into the *Sal*I and *Not*I sites of the *Rsr*II-modified pPC86. *Eco*RI and *Spe*I sites were then introduced within the *trxA* *Rsr*II site. The oligonucleotides encoding the peptide aptamers were cloned into the *Eco*RI and *Spe*I sites of the resulting vector.

TABLE 13

Contig	ATCC No.	SEQ ID NO.	Length (base pairs)
b527d12-h contig302G	980720	5	3096
b527d12-h contig306G	980720	6	26928
b527d12-h contig307G	980720	7	29430
b527d12-h contig308G	980720	8	33769
b527d12-h contig309G	980720	9	72049
b200e21-h contig1	980812	10	8705
b200e21-h contig4	980812	11	66933

Example 4

Generation of antibodies

In each of the following antibody-generating examples, the synthesis of these linear peptides is followed by injection into two New Zealand Rabbits. Subsequent boosts and

bleeds are taken according to a standard ten-week protocol. The end-user receives back 5 mgs of peptide, aliquots of pre-bleeds, roughly 80 ml of crude sera from each of the two rabbits and, and ELISA titration data is obtained.

Generation of LRP5 Polymorphism-specific antibodies. Antibodies were generated to the following peptides to obtain antibodies which distinguish the HBM polymorphism versus wild-type LRP5/Zmax: MYWTDWVETPRIE (SEQ ID NO:) (mutant peptide) and MYWTDWGETPRIE (SEQ ID NO:) (wild-type peptide for negative selection). Immunofluorescence data confirmed that the antibody, after affinity purification, is specific for HBM and does not recognize LRP5 (Fig. 32).

Generation of LRP5 Monospecific antibodies. LRP5 monospecific polyclonal antibodies were generated to the following amino acid sequences of LRP5: Peptide 1 (a.a. 265-277) - KRTGGKRKEILSA (SEQ ID NO:), Peptide 2 (a.a. 1178-1194) - ERVEKTTGDKRTRIQGR (SEQ ID NO:), and Peptide 3 (a.a. 1352-1375) - KQQCDSFPDCIDGSDE (SEQ ID NO:). Immunofluorescence confirmed that the antibody generated detects LRP5.

Generation of Dkk-1 monospecific polyclonal antibodies. Dkk-1 monospecific polyclonal antibodies were generated to the following amino acid sequences of Dkk-1: Peptide 1 - GNKYQTIDNYQPYPYPC (SEQ ID NO:), Peptide 2- LDGYSRRTTLSSKMYHTKGQEG (SEQ ID NO:), Peptide 3 - RIQKDHHQASNSSLHTCQRH (SEQ ID NO:), Peptide 4 - RGEIETITESFGND (SEQ ID NO:), and Peptide 5 - EIFQRCGEGLSCRIQKD (SEQ ID NO:) of human Dkk-1. Western Blots demonstrated that the antibodies generated against peptides 2 and 4 are specific toward Dkk-1.

Example 5

Effects of exogenous Dkk-1 on Wnt-mediated signaling in the *Xenopus* embryo assay

Xenopus embryos are an informative and well-established *in vivo* assay system to evaluate the modulation of Wnt signaling (McMahon *et al.*, 1989 *Cell* 58: 1075-84; Smith *et*

al., 1991 reviewed in Wodarz and Nusse, 1998).

Modification of the Wnt signaling pathway can be visualized by examining the embryos for a dorsalization phenotype (duplicated body axis) after RNA injection into the ventral blastomere at the 4- or 8-cell stage. On the molecular level, phenotypes can be analyzed by looking for expression of various marker genes in stage 10.5 embryos. Such markers would include general endoderm, mesoderm, and ectoderm markers as well as a variety of tissue-specific transcripts.

Analysis can be done by RT-PCR/TaqMan® and can be done on whole embryo tissue or in a more restricted fashion (microdissection). Because this system is very flexible and rapid, by injecting combinations of transcripts, such as HBM and different Wnts or Wnt antagonists, the mechanism of HBM in the Wnt pathway can thereby be dissected. Furthermore, investigations are conducted to determine whether Zmax/LRP5 and HBM differentially modulate Wnt signaling either alone, or in combination with other components. Previous studies have demonstrated that LRP6 alone or LRP5 + Wnt5a were able to induce axis duplication (dorsalization) in this system (Tamai *et al.*, 2000 *Nature* 407: 530-35).

Constructs for Xenopus Expression (Vector pCS2⁺). Constructs were prepared using the vector pCS2⁺. DNA inserts were subcloned in the sense orientation with respect to the vector SP6 promoter. The pCS2⁺ vector contains an SV40 virus polyadenylation signal and T3 promoter sequence (for generation of antisense mRNA) downstream of the insert.

Full length Zmax/LRP5 and HBM ORF cDNA: Insert cDNA was isolated from the full length cDNA retrovirus constructs (with optimized Kozak sequences) by *Bgl*II-*Eco*RI digestion and subcloned into the *Bam*HI-*Eco*RI sites of the pCS2⁺ vector. cDNAs encoding a HBM-like molecule could be subcloned into pCS2⁺ vectors and processed similarly by one of ordinary skill.

Full length XWnt8: This cDNA was PCR amplified from a *Xenopus* embryo cDNA library using oligos 114484 (SEQ ID NO:) (5'-CAGTGAATTCACCATGCAAAACACC ACTTTGTTC-3') and 114487 (SEQ ID NO:) (5'-CAGTTGCGGCCGCTCATCTCCGGTG GCCTCTG-3'). The oligos were designed to amplify the ORF with a consensus Kozak sequence at the 5' end as determined from GenBank #X57234. PCR was carried out using

the following conditions: 96°C, 45 sec.; 63°C, 45 sec.; 72°C, 2 min. for 30 cycles. The resulting PCR product was purified, subcloned into pCRII-TOPO (Invitrogen Corp.), sequence verified, and digested with *Bam*HI/*Xho*I. This insert was subcloned into the vector at the *Bam*HI-*Xho*I sites.

5 Full length Wnt5a: A murine Wnt5a cDNA clone was purchased from Upstate Biotechnology (Lake Placid, NY) and subcloned into the *Eco*RI site of the vector. Sequencing confirmed insert orientation.

10 Full length human Dkk-1: A human cDNA with GenBank accession number AF127563 was available in the public database. Using this sequence, PCR primers were designed to amplify the open reading frame with a consensus Kozak sequence immediately upstream of the initiating ATG. Oligos 117162 (SEQ ID NO:) (5'-CAATAGTCGACGAATTCACCATGGCTCTGGGCGCAGCGG-3') and 117163 (SEQ ID NO:) (5'-GTATTGCGGCCGCTCTAGATTAGTGTCTCTGACAAGTGTGAA-3') were used to screen a human uterus cDNA library by PCR. The resulting PCR product was
15 purified, subcloned into pCRII-TOPO (Invitrogen Corp.), sequence verified, and digested with *Eco*RI/*Xho*I. This insert was subcloned into the pCS2⁺ vector at the *Eco*RI-*Xho*I sites.

20 Full length human Dkk-2: A full length cDNA encoding human Dkk-2 was isolated to investigate the specificity of the Zmax/LRP5/HBM interaction with the Dkk family of molecules. Dkk-1 was identified in yeast as a potential binding partner of Zmax/LRP5/HBM. Dkk-1 has also been shown in the literature to be an antagonist of the Wnt signaling pathway, while Dkk-2 is not (Krupnik *et al.*, 1999). The Dkk-2 full length cDNA serves as a tool to discriminate the specificity and biological significance of Zmax/LRP5/HBM interactions with the Dkk family (*e.g.*, Dkk-1, Dkk-2, Dkk-3, Dkk-4, Soggy, their homologs and variant, etc.). A human cDNA sequence for Dkk-2 (GenBank
25 Accession No. NM_014421) was available in the public database. Using this sequence, PCR primers were designed to amplify the open reading frame with a consensus Kozak sequence immediately upstream of the initiating ATG. Oligos 51409 (SEQ ID NO:) (5'-CTAACGGATCCACCATGGCCGCGTTGATGCGG-3') and 51411 (SEQ ID NO:) (5'-GATTCGAATTCTCAAATTTCTGACACACATGG-3') were used to screen human

embryo and brain cDNA libraries by PCR. The resulting PCR product was purified, subcloned into pCRII-TOPO, sequence verified, and digested with *Bam*HI/*Eco*RI. This insert was subcloned into the pCS2⁺ vector at the *Bam*HI-*Eco*RI sites.

Full length LRP6 was isolated from the pED6dpc4 vector by *Xho*I-*Xba*I digestion. The full length cDNA was reassembled into the *Xho*I-*Xba*I sites of pCS2⁺. Insert orientation was confirmed by DNA sequencing.

mRNA Synthesis and Microinjection Protocol. mRNA for microinjection into *Xenopus* embryos is generated by *in vitro* transcription using the cDNA constructs in the pCS2⁺ vector described above as template. RNA is synthesized using the Ambion mMessage mMachine high yield capped RNA transcription kit (Cat. #1340) following the manufacturer's specifications for the Sp6 polymerase reactions. RNA products were brought up to a final volume of 50 µl in sterile, glass-distilled water and purified over Quick Spin Columns for Radiolabeled RNA Purification G50-Sephadex (Roche, Cat. #1274015) following the manufacturer's specifications. The resulting eluate was finally extracted with phenol:chloroform:isoamyl alcohol and isopropanol precipitated using standard protocols (Sambrook *et al.*, 1989). Final RNA volumes were approximately 50 µl. RNA concentration was determined by absorbance values at 260 nm and 280 nm. RNA integrity was visualized by ethidium bromide staining of denaturing (formaldehyde) agarose gel electrophoresis (Sambrook *et al.*, 1989). Various amounts of RNA (2 pg to 1 ng) are injected into the ventral blastomere of the 4- or 8-cell *Xenopus* embryo. These protocols are described in Moon *et al.*, 1989 *Technique-J. of Methods in Cell & Mol. Biol.* 1: 76-89; and Peng, 1991 *Meth. Cell. Biol.* 36: 657-62.

Screening for Duplicated Body Axis. *In vitro* transcribed RNA is purified and injected into a ventral blasomere of the 4- or 8-cell *Xenopus* embryo (approx. 2 hours post-fertilization). At stage 10.5 (approx. 11 hours post-fertilization), the injected embryos are cultured for a total of 72 hours and then screened for the presence of a duplicated body axis (dorsalization) (Fig. 33). Using XWnt8-injected (2-10 pg) as a positive control (Christian *et al.*, 1991) and water-injected or non-injected embryos as negative controls, we replicated the published observation that Zmax(LRP5) + Wnt5a (500 and 20 pg, respectively) could induce

axis duplication. Wnt5a (20 pg) alone could not induce axis duplication (as previously reported by Moon *et al.*, 1993). We have also injected GFP RNA (100-770 pg) as a negative control to show that the amount of RNA injected is not perturbing embryo development (not shown). Strikingly, HBM + Wnt5a (500 and 20 pg, respectively) yielded an approximately
 5 3.5 fold more robust response of the phenotype ($p=0.043$ by Fisher's exact test) compared to Zmax(LRP5) + Wnt5a, suggesting that the HBM mutation is activating the Wnt pathway (Figs. 34 and 35). The HBM/Wnt5a embryos also appear to be more "anteriorized" than the Zmax(LRP5)/Wnt5a embryos, again suggestive of a gain-of-function mutation.

The role of Dkk-1 as a modulator of Zmax/LRP5- and HBM-mediated Wnt signaling
 10 was investigated. Literature reports have previously characterized *Xenopus* and murine Dkk-1 as antagonists of the canonical Wnt pathway in the *Xenopus* system (Glinka *et al.*, 1998 *Nature* 391: 357-62). Using the human Dkk-1 construct, a dose-response assay was performed to confirm that our construct was functional and to identify the optimal amount of RNA for microinjection. Using 250 pg/embryo of hDkk-1 RNA, over 90% ($p<0.001$) of the
 15 embryos were observed to display enlarged anterior structures (big heads) as anticipated from the published reports (Fig. 36).

The mechanism of hDkk-1 modulation of Wnt signaling in the presence of Zmax/LRP5 or HBM was also investigated. Without any hDkk-1 present, it was confirmed that HBM + Wnt5a was a more potent activator of Wnt signaling than Zmax/LRP5 + Wnt5a
 20 ($p<0.05$). Interestingly, in the presence of hDkk-1 (250 pg), Zmax/LRP5-mediated Wnt signaling was repressed ($p<0.05$) but hDkk-1 was unable to repress HBM-mediated Wnt signaling ($p<0.01$) (Fig. 37). The specificity of this observation can be further addressed by investigating other members of the Dkk family, other Wnt genes, LRP6, additional Zmax/LRP5 mutants, and the peptide aptamers.

Example 6

Effects of exogenous Dkk and LRP5 on Wnt signaling in the TCF-luciferase Assay

Wnt activity can be antagonized by many proteins including secreted Frizzled related proteins (SFRPs), Cerberus, Wnt Inhibitory Factor-1 and Dkk-1 (Krupnik *et al.*, 1999). The

Dkk family of proteins consists of Dkk-1-4 and Soggy, a Dkk-3-like protein. Dkk-1 and Dkk-4 have been shown to antagonize Wnt mediated *Xenopus* embryo development, whereas Dkk-2, Dkk-3, and Soggy do not. Unlike many of these proteins that antagonize Wnt activity by directly interacting with Wnt proteins, Dkk-1 acts by binding to two recently identified
5 Wnt coreceptors, LRP5 and LRP6 (Mao *et al.*, 2001; Bafico *et al.*, 2001). The details of this interaction have been examined by the present inventors and Mao *et al.* using deletion constructs of LRP6, which demonstrated that EGF repeats 3 and 4 are important for Dkk-1 interaction. Accordingly, the activity of two Dkk proteins, Dkk-1 and Dkk-2, were investigated with various Wnt members, LRP5, LRP6, and the mutant form of LRP5,
10 designated HBM. The present invention explores whether there is any functional difference between LRP5 and HBM with regard to Dkk action on Wnt mediated signaling. Various reagents were developed, including Dkk-1 peptides, constrained LRP5 peptide aptamers, constrained Dkk-1 peptide aptamers and polyclonal antibodies to Dkk-1 (in Example 4 above) to identify factors that mimic HBM mediated Wnt signaling.

15 **Methods.** Various LRP5 constrained peptides were developed. Specifically, four peptides that interact with the LBD of LRP5 (Fig. 38, constructs OST259-262 in Fig. 39) and three peptides that interact with the cytoplasmic domain of LRP5 (constructs OST266-OST268 in Fig. 39). In addition two Dkk-1 peptides were developed: constructs OST264 and OST265 in Fig. 39, corresponding to Dkk-1 amino acids 139-266 and 96-245, containing
20 the smallest region of Dkk-1 that interacts with LRP5 (Fig. 40). The cDNA clones encoding the LRP5 LBD interacting peptides and the Dkk-1 peptides were subcloned into pcDNA3.1 with the addition of a Kozak and signal sequence to target the peptide for secretion. The constructs encoding the three peptides interacting with the cytoplasmic domain of LRP5 were also subcloned into pcDNA3.1. However, these latter constructs do not contain a signal
25 sequence.

HOB-03-CE6 osteoblastic cells developed by Wyeth Ayerst (Philadelphia, PA) were seeded into 24-well plates at 150,000 cells per well in 1 ml of the growth media (D-MEM/F12 phenol red-free) containing 10% (v/v) heat-inactivated FBS, 1X penicillin streptomycin, and 1X Glutamax-1, and incubated overnight at 34°C. The following day, the

cells were transfected using Lipofectamine 2000® (as described by the manufacturer, Invitrogen) in OptiMEM (Invitrogen) with 0.35 μg /well of LRP5, HBM, or control plasmid DNA (empty vector pcDNA3.1) and either Wnt1 or Wnt3a plasmid DNA. Similar experiments were performed with LRP6 plasmid DNA (0.35 μg /well) or a control pEDdpc4 empty vector. Furthermore, each of these groups were then divided into three groups, those receiving 0.35 μg /well Dkk-1, Dkk-2, or pcDNA3.1 control DNA. All wells were transfected with 0.025 μg /well of CMV beta-galactosidase plasmid DNA and 0.35 μg /well 16X TCF(AS)-luciferase reporter DNA (developed by Ramesh Bhat, Wyeth-Ayerst (Philadelphia, PA)). After 4 hours of incubation, the cells were rinsed and 1 ml of fresh growth media was added to each well. The cells were cultured overnight at 34°C, followed by a wash and a change of media. Cells were cultured for an additional 18-24 hours at 37°C. Cells were then lysed with 50 μl /well of 1X lysis buffer. The extracts were assayed for beta-galactosidase activity (Galacto Reaction Buffer Diluent & Light Emission Accelerator, Tropix) using 5 μl extract + 50 μl beta-galactosidase diluent and luciferase activity (Luciferase Assay Reagent, Promega) using 20 μl extract.

U2OS human osteosarcoma cells were also utilized. U2OS cells (ATCC) were seeded into 96-well plates at 30,000 cells per well in 200 μl of the growth media (McCoy's 5A) containing 10% (v/v) heat-inactivated FBS, 1X penicillin streptomycin, and 1X Glutamax-1, and incubated overnight at 37°C. The following day, the media was replaced with OptiMEM (Invitrogen) and cells were transfected using Lipofectamine 2000® (as described by the manufacturer, Invitrogen) with 0.005 μg /well of LRP5, HBM, LRP6 or control plasmid DNA (empty vector pcDNA3.1) and either Wnt1 (0.0025 μg /well) or Wnt3a (0.0025 μg /well) plasmid DNA. In addition, the 16x-(AS) TCF-TK-firefly-luciferase and control TK-renilla luciferase (Promega Corp.) were co-transfected at 0.3 μg /well and 0.06 μg /well respectively in all experiments. Furthermore, each of these groups was then divided into different groups, those receiving 0.05 μg /well Dkk-1, Dkk-2, Dkk3, Dkk1-Alkaline Phosphatase (AP), mutant Dkk-1 (C220A), Soggy or pcDNA3.1 control DNA. In other experiments, cells were co-transfected with 0.005 μg /well of LRP5, 0.0025 μg /well of Wnt1 or Wnt3a (using 0.0025 μg /well of a control pcDNA3.1) with LRP5-interacting aptamers

(0.05 μ g/well). Cells were cultured for an additional 18-20 hours at 37°C. Culture medium was removed. Cells were cultured for an additional 18-20 hours at 37°C. Culture medium was removed. Cells were then lysed with 100 μ l/well of 1X Passive Lysis Buffer (PLB) of Dual Luciferase Reagent kit (DLR-kit-Promega Corp.) 20 μ l of the lysates were combined with LARII reagent of DLR-kit and assayed for TCF-firefly luciferase signal in Top Count (Packard) instrument. After measuring the Firefly readings, 100 μ l of the "Stop and Glo" reagent of DLR kit that contains a quencher and a substrate for renilla luciferase was added into each well. Immediately the renilla luciferase reading was measured using the Top Count (Packard) Instrument. The ratios of the TCF-firefly luciferase to control renilla readings were calculated for each well and the mean ratio of triplicate or more wells was expressed in all data.

Results. The results of these experiments demonstrate that Dkk-1, in the presence of Wnt1 and LRP5, significantly antagonized TCF-luciferase activity (Fig. 41). In marked contrast, Dkk-1 had no effect on HBM/Wnt1 mediated TCF-luciferase activity (Fig. 41). In similar experiments, Dkk-1 was also able to antagonize LRP5/Wnt3a but not HBM/Wnt3a mediated TCF-luciferase activity (Fig. 42). These results indicate that the HBM mutation renders Dkk-1 inactive as an antagonist of Wnt1 and Wnt3a signaling in HOB03CE6 osteoblastic cells. In other experiments with Wnt1, Dkk-1 had no effect on LRP5 or HBM mediated TCF-luciferase activity (Fig. 41). In contrast, with either LRP5 or HBM in the presence of Wnt3a, Dkk-2 was able to antagonize the TCF-luciferase activity (Fig. 42). These latter results indicate that the HBM mutation has no effect on Dkk-2 action in the presence of Wnt3a. Experiments were also performed using the closely related LRP6 cDNA in HOB-03-CE6 cells. In these experiments, LRP6/Wnt1 and LRP6/Wnt3a mediated TCF-luciferase were regulated in the same manner as LRP5. Specifically, Dkk-1 antagonized LRP6/Wnt1 mediated TCF-luciferase activity, whereas Dkk-2 had no effect (Fig. 41). However, similar to the action of Dkk-2 with LRP5/Wnt3a, Dkk-2 was able to antagonize LRP6/Wnt3a mediated TCF-luciferase activity (Fig. 42).

The results in the U2OS cells show a robust effect of the OST262 LRP5 peptide aptamer activation of Wnt signaling in the presence of Wnt3a (Fig. 43). These functional

results are confirmed by the results shown below in Example 7 using LRP5 peptide aptamers in the *Xenopus* assay. Such results affirmatively demonstrate that the effects of small molecules on LRP5/LRP6/HBM signaling can be detected using the TCF-luciferase assay.

These data demonstrate that there is a functional difference between LRP5 and HBM regarding the ability of Dkk-1 to antagonize Wnt1 and Wnt3a signaling. These data and previous data showing that Dkk-1 directly interacts with LRP5 suggests that the inability of Dkk-1 to antagonize HBM/Wnt signaling may in part contribute to the HBM phenotype. These experiments further demonstrate the ability to test various molecules (e.g., small molecules, aptamers, peptides, antibodies, LRP5 interacting proteins or Dkk-1 interacting proteins, and the like) for a LRP5 ligand that mimics HBM mediated Wnt signaling or factors that block Dkk-1 interaction with LRP5.

This was the assay that was used to show the responsiveness of two HBM-like variants. See Figs. 29 and 30. The data also demonstrates that these variants are less susceptible to modulation by Dkk.

Example 7

Cell-Based Functional High-Throughput Assay

To develop a high throughput assay, the TCF-luciferase assay described in Example 6 was modified utilizing low level expression of endogenous LRP5/6 in U2OS and HEK293 cells. However, HOB-03-CE6 cells and any other cells which show a differential response to Dkk depending on whether LRP5, LRP6 or HBM are expressed. Using U2OS (human osteosarcoma) and HEK293 (ATCC) cells, the TCF-luciferase and tk-Renilla reporter element constructs were co-transfected along with Wnt3a/1 and Dkk. Wnt3a alone, by using endogenous LRP5/6, was able to stimulate TCF reporter gene activation. When Dkk, is co-transfected with Wnt3a/Wnt 1 and reporters (TCF-luci and tk-Renilla), Dkk represses reporter element activity. In addition, the TCF-luci signal is activated by Wnt3a/Wnt1 can be repressed by the addition of Dkk-enriched conditioned media to the cells containing Wnt3a/Wnt1 and reporters. The assay is further validated by the lack of TCF-reporter inhibition by a point mutant construct (C220A) of Dkk1.

The Dkk-mediated repression of the reporter is dependent upon the concentration of transfected Dkk cDNA or on the amount of Dkk-conditioned media added. In addition, the Dkk-mediated reporter suppression can be altered by the co-transfection of LRP5, LRP6, and HBM cDNAs in the U2OS or HEK293 cells. In general, U2OS cells show greater sensitivity to Dkk-mediated reporter suppression than that in HEK-293 cells. In U2OS cells, the transfection of LRP5/LRP6/HBM/HBM-like cDNA leads to moderate activation of TCF-luciferase in the absence of Wnt3a/Wnt1 transfection. This activation presumably utilizes the endogenous Wnts present in U2OS cells. Under this condition, Dkk1 can repress TCF-luciferase and shows a differential signal between LRP5 and HBM. By co-transfecting Wnt3a/Wnt1, there is a generalized increase in the TCF-luciferase signal in the assay. Further, one can detect Dkk-mediated differential repression of the reporter due to LRP5 and HBM cDNA expression as well as between LRP5 and LRP6 cDNA. The repression is maximal with LRP6, moderate with LRP5, and least with HBM cDNA expression. In addition, the assay can detect the functional impact of the LRP5 interacting peptide aptamers (Fig. 38), Dkk1 interacting aptamers and binding domains of Dkk-1 (Fig. 40; OST264 and OST265 of Figs. 39 and 44).

Using this system with a suppressed Wnt-TCF signal due to the presence of both Dkk and Wnt3a, one can screen for compounds that could alter Dkk modulation of Wnt signaling, by looking for compounds that activate or the TCF-luciferase reporter, and thereby relieve the Dkk-mediated repression of the Wnt pathway. Such compounds identified may potentially serve as HBM-mimetics and be useful, for example, as osteogenic therapeutics. Data generated from this high throughput screen are demonstrated in Figs. 45-47. Fig. 45 shows that Dkk1 represses Wnt3a-mediated signaling in U2OS bone cells. Fig. 46 demonstrates the functional differences between LRP5, LRP6, and HBM. Dkk-1 represses LRP6 and LRP5 but has little or no effect on HBM-generated Wnt1 signaling in U2OS cells. Fig. 47 demonstrates the differential effects of various Dkk family members and modified Dkks, including Dkk-1, a mutated Dkk-1 (C220A), Dkk-1-AP (modified with alkaline phosphatase), Dkk-3, and Soggy.

Example 8

DKK/LRP5/6/HBM/HBM-like ELISA Assay

A further method to investigate Dkk binding to LRP or HBM and HBM like polypeptides is via ELISA assay. Two possible permutations of this assay are exemplified.

5 LRP5 is immobilized to a solid surface, such as a tissue culture plate well. One skilled in the art will recognize that other supports such as a nylon or nitrocellulose membrane, a silicon chip, a glass slide, beads, etc. can be utilized. In this example, the form of LRP5 used is actually a fusion protein where the extracellular domain of LRP5 is fused to the Fc portion of human IgG. The LRP5-Fc fusion protein is produced in CHO cell extracts from stable cell

10 lines. The LRP5-Fc fusion protein is immobilized on the solid surface via anti-human Fc antibody or by Protein-A or Protein G-coated plates, for example. The plate is then washed to remove any non-bound protein. Conditioned media containing secreted Dkk protein or secreted Dkk-epitope tagged protein (or purified Dkk or purified Dkk-epitope tagged protein) is incubated in the wells and binding of Dkk to LRP is investigated using antibodies to either

15 Dkk or to an epitope tag. Dkk-V5 epitope tagged protein would be detected using an alkaline phosphatase tagged anti-V5 antibody.

Alternatively, the Dkk protein could be directly fused to a detection marker, such as alkaline phosphatase. Here the detection of the Dkk-LRP interaction can be directly investigated without subsequent antibody-based experiments. The bound Dkk is detected in

20 an alkaline phosphatase assay. If the Dkk-alkaline phosphatase fusion protein is bound to the immobilized LRP5, alkaline phosphatase activity would be detected in a colorimetric readout. As a result, one can assay the ability of small molecule compounds to alter the binding of Dkk to LRP using this system. Compounds, when added with Dkk (or epitope-tagged Dkk) to each well of the plate, can be scored for their ability to modulate the

25 interaction between Dkk and LRP based on the signal intensity of bound Dkk present in the well after a suitable incubation time and washing. The assay can be calibrated by doing cold competition experiments with unlabeled Dkk or with a second type of epitope-tagged Dkk. Any small molecule that is able to modulate the Dkk-LRP interaction may be a suitable therapeutic candidate, more preferably an osteogenic therapeutic candidate.

Example 9

Functional Evaluation of Peptide Aptamers in *Xenopus*

The constrained peptide aptamers constructs OST258-263 (where 258 contains the signal sequence by itself and 263 contains an irrelevant constrained peptide) (Figs. 39 and 44) were used to generate RNA substantially as described in Example 6, except the vector was linearized by restriction endonuclease digestion and RNA was generated using T7 RNA polymerase.

Aptamer RNA was injected at 250 pg per blastomere using the protocol of Example 6. Wnt signaling was activated, as visualized by embryo dorsalization (duplicated body axis) with aptamers 261 and, more strongly, 262. The results of this assay are shown in Figs. 48 and 49. These results suggest that aptamers 261 and 262 are able to activate Wnt signaling possibly by binding to the LBD of LRP, thereby preventing the modulation of LRP-mediated signaling by Dkk.

The aptamers of the present invention can serve as HBM-mimetics. In the *Xenopus* system they are able to induce Wnt signaling all by themselves. They may also serve as tools for rational drug design by enhancing the understanding of how peptides are able to interact with LRP and modulate Wnt signaling at the specific amino acid level. Thus, one would be able to design small molecules to mimic their effects as therapeutics. In addition, the aptamers identified as positives in this assay may be used as therapeutic molecules themselves.

Example 10

Homogenous Assay

An excellent method to investigate perturbations in protein-protein interactions is via Fluorescence Resonance Energy Transfer (FRET). FRET is a quantum mechanical process where a fluorescent molecule, the donor, transfers energy to an acceptor chromophore molecule which is in close proximity. This system has been successfully used in the literature to characterize the intermolecular interactions between LRP5 and Axin (Mao *et al.*, *Molec. Cell Biol.* 7: 801-9). There are many different fluorescent tags available for such

studies and there are several ways to fluorescently tag the proteins of interest. For example, CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein) can be used as donor and acceptor, respectively. Fusion proteins, with a donor and an acceptor, can be engineered, expressed, and purified.

5 For instance, purified LRP protein, or portions or domains thereof, fused to CFP and purified Dkk protein, or portions or domains thereof that interact with Dkk or LRP respectively, fused to YFP can be generated and purified using standard approaches. If LRP-CFP and Dkk-YFP are in close proximity, the transfer of energy from CFP to YFP will result in a reduction of CFP emission and an increase in YFP emission. Energy is supplied with an
10 excitation wavelength of 450 nm and the energy transfer is recorded at emission wavelengths of 480 nm and 570 nm. The ratio of YFP emission to CFP emission provides a gauge for changes in the interaction between LRP and Dkk. This system is amenable for screening small molecule compounds that may alter the Dkk-LRP protein-protein interaction. Compounds that disrupt the interaction would be identified by a decrease in the ratio of YFP
15 emission to CFP emission. Such compounds that modulate the LRP-Dkk interaction would then be considered candidate HBM mimetic molecules. Further characterization of the compounds can be done using the TCF-luciferase or *Xenopus* embryo assays to elucidate the effects of the compounds on Wnt signaling.

20 While the above example describes a cell-free, solution-phase assay using purified components, a similar cell-based assay could also be performed. For example, LRP-CFP fusion protein can be expressed in cells. The Dkk-YFP fusion protein then could be added to the cells either as purified protein or as conditioned media. The interaction of LRP and Dkk is then monitored as described above.

25 **Example 11**

Identification of Variants of LRP5/Zmax1/HBM

Because the YWTD repeats constitute the major part of domain which interacts with Dkk-1, a search was begun from the repeats to look for protein folds that contain such repeats. Springer, 1998 *J. Mol. Biol.* 283: 837-62 proposed that these YWTD repeats, which

had been previously described as "spacers" and considered to have no defined structure, are in fact the propeller-blade subdomains of highly structured six-bladed β -propellers. Springer (1998) therefore proposed two theoretically modeled protein structures with such YWTD repeats.

5 **The Model.** A set of LRP5 propeller domain sequences from mouse and human were assembled and aligned using CLUSTALW, together with the sequences corresponding to two of the modeled YWTD-propeller structures of Springer (1998) [1lpx, based on the LRP1 β -propeller domain no. 7 from chicken (SwissProt LRP1_CH7), and 1ndx, based on human nidogen (Swiss-Prot NIDO_HU1)]. The alignment was manually edited based on the
10 rationale and knowledge of protein structures. Fig. 50 depicts a schematic model of LRP5. The secondary structures consist exclusively of β -strands and turns. The secondary structure assignments were verified by the predictive program, DSC (Discrimination of Protein Secondary Structure Class available at <http://bioweb.pasteur.fr/seqanal/interfaces/dsc.html>). Preliminary checks of the exon-intron boundaries were performed manually. A tertiary-
15 structure homology model for human LRP5 using the 1lpx model as a template was built with InsightII (Accelrys Inc., San Diego, CA) and examined using the graphics display programs from Insight II and Rasmol (freely available at <http://www.umass.edu/microbio/rasmol/index2.html>).

20 **Results.** Based on the homology alignments and modeling, the domain diagram of LRP5 in Fig. 50 was obtained. Fig. 51 shows the complete alignment and secondary structure assignments for the propeller domains of mouse LRP5 and LRP6, human LRP5 and human LRP6, and the sequences corresponding to the two theoretical models constructed by Springer (1998).

25 To a good first approximation, the sequence alignments support the proposal that there are four 6-bladed β -propeller domains in human LRP5. This model nicely accommodates the sequence (as should be expected since it derives from a carefully constructed model of a chicken LRP propeller domain). The overall shape of this domain resembles a disk with inward-sloping sides and a hole down the middle. The polypeptide chain enters and leaves from the same (comparatively flat) "bottom" surface. As indicated in

the alignment (Fig. 51) and in several of the structure illustrations, the G171V mutation falls into a loop on the outer or “top” face of the domain. This immediately suggests that – contrary to what one might expect, given the role of highly conserved glycine residues in determining protein folds – the mutation should not have any significant effect on the domain’s tertiary structural stability. It is rather more likely that the mutation interferes with binding of some ligand, possibly a macromolecule (Smith *et al.*, 1999 *Trends Biochem. Sci.* 24: 181-5) or alters a protein-protein interaction.

Closer scrutiny of the protein in the vicinity of G171 shows that this residue sits at the bottom of a pocket on the outer edge of the top surface of the domain. Moreover, one side of this pocket consists of a cluster of very hydrophobic side chains, while the other side has more polar groups, most notably glutamate 172. In the mutant form of the protein, with V171 in place of the glycine, the pocket largely disappears as the isopropyl group of valine replaces the hydrogen side chain of glycine. It is easy to understand how such a substitution could seriously disrupt ligand binding if the ligand normally protrudes into the pocket. By the same rationale, any other mutations that block such pockets in any of the four propellers could also result in impaired ligand binding.

The reasonableness of the proposed model is based on Springer's results as well as additional data. For example, data derived from the typical features of such protein domains (including the well-studied YWTD repeats), e.g., that are robust and rigid tertiary structures. Loop conformations may vary somewhat, especially for the longer loops, but the basic features of the protein scaffold are almost certainly well predicted. Of course we would be on somewhat more secure ground if there were a crystal structure for one of these domains. That naturally implies that some of the more interesting features, such as the identity of residues exposed on the outer, putative ligand-binding surface, will also be well-predicted. Note that in Fig. 51 these residues are marked in red on the sequences corresponding to Springer’s (1998) models. Even though different models could possibly be built that are not dependent on Springer’s assumptions and lead to a different topology of the propeller, this would not alter the conclusion that the G171V mutation lies in a surface loop.

At the very least, the model affords an opportunity to think about this protein in

molecular terms and that should facilitate both experimental design and evaluation of results, *i.e.*, candidate epitopes might be selected more rationally. One interesting aspect, evoked forcefully by the diagram in Fig. 50, is the question of how the different modular propeller domains interact. The EGF-like domains are well-studied (dozens of crystal and NMR structures exist, see for example Bork *et al.*, 1996 *Quart. Rev. Biophys.* 29: 119-67). The EGF-like domains consist mainly of a couple of β -hairpins, cross-linked by disulfide bonds, and in some cases they have tightly bound calcium ions, which must further stabilize their structure. Interestingly, in contrast to the YWTD propellers, the EGF-like domains have the polypeptide chain entering at one end and exiting at the other. Thus, mechanically, these domains could act as spacers, even swivels (since rotation can occur around the tails that extend from either end). Given the size of the propellers, there are almost certainly inter-domain interactions between them as they assemble in some higher-order structure, connected by the EGF-like domains. Because of the topology of the propellers, their "top" faces would have to face outward since the EGF-like domains will tie them together by their protruding N- and C-terminal extensions.

A more sophisticated analysis of the β -propeller structure was modeled using X-ray crystallographic data from the LDL receptor. The primary amino acid sequences of the various beta propeller domains of LRP5 (and other domains) were used to develop homology models of their 3-dimensional structures according to the following method:

(A) Search for suitable structural templates and check sequence identity with target. The ExNRL-3D database (derived from the Protein Data Bank (<http://pdb.ccdc.cam.ac.uk/pdb/>) was searched using BLASTP2 (Altschul *et al.*, 1990 *J. Mol. Biol.* 215: 403-10; Huang *et al.*, 1991 *Adv. Appl. Math.* 12: 337-67; and Peitsch, 1995 *PDB Quart. Newsletter* 72: 4) to find all similarities of the target sequence with sequences of known structure. This program selects all templates with sequence identities above 25% and projected model size greater than 20 residues. This step also detects domains that can be modelled based on unrelated templates. Use of this step resulted in the selection of Protein Data Bank (PDB) structure, 1IJQ ("Crystal Structure Of The LDL Receptor YWTD-EGF Domain Pair") (Jeon *et al.*,

2001 *Nat. Struct. Biol.* 8: 499).

(B) Create ProModII (Altschul *et al.*, 1990; Huang *et al.*, 1991; and Peitsch, 1995) jobs and generate models with ProModII.

(C) Superimpose related 3D structures. This method is based upon the diagonals of sequence similarity (the Dynamic sequence alignment algorithm (SIM)(Altschul *et al.*, 1990; Huang *et al.*, 1991; and Peitsch, 1995). Specifically, the following steps are performed:

(i) Primary match: Regions with sequence similarity are selected automatically (manual selection is also possible) and the corresponding residues matched in three-space. (ii) The primary match is further refined using expanding context spheres.

(D) Generate a multiple alignment comprising the sequences to be modeled.

(E) Generate a framework for the new sequence.

(F) Based on the topological arrangement of corresponding atoms: (a) atoms which occupy a similar portion of space and are expected to have a structural counterpart in the new structure are used to compute the framework coordinates (averaged positions); (b) Side chains with fully incorrect geometries are removed.

(G) Rebuild the lacking loops based on the geometry of the loop stems: (a) the stems of the loops to rebuild are used to scan a database of structural fragments derived from the Brookhaven Data Bank. Either the best fitting fragment or a framework derived from the five best fragments, is used as the new loop, or (b) the conformational space is searched using a CSP approach (seven allowed Φ - Ψ angle combinations, space allocation for the loop, space allocation for each α -carbon in the loop). Both methods will place only α -carbons when loops are in steric conflict with the surrounding context.

(H) Rebuilt incomplete backbones based on the position of the α -carbons, the backbone was rebuilt using a set of seven allowed Φ - Ψ angle combinations and a database of backbone fragments (a sliding window of five residues was run through the protein sequence. The best matching backbone fragment for each overlapping pentapeptide was then stored. A framework for the main chain atoms is derived from

these peptides, using only the coordinates of the three central residues of each pentapeptide.

(I) The model's structural quality is verified and packing is checked as follows:

(i) Verification of structural quality is based on the method described by R. Luthy *et al.*, 1992 *Nature* 356: 83-85. This method analyzes the 3D context of each residue and allows the identification of mis-folded regions.

(ii) Packing is checked based on a probe accessible surface (Connolly surface) computation and a cubic grid passed through the structure. Inside and outside surfaces are detected and the center and size of each cavity is computed. The algorithm then compares the size and distribution of these cavities between the model and the structure in order to detect possibly mis-folded regions.

(J) Refine structure by energy minimization and molecular dynamics based on force field computations following Gromos96 (*Methods for the evaluation of long-range electrostatic forces in computer simulations of molecular systems*, IN

COMPUTER SIMULATION OF BIOMOLECULAR SYSTEMS, THEORETICAL AND EXPERIMENTAL APPLICATIONS, Vol. 2, 182-212 (W.F. van Gunsteren *et al.*, eds., Escom Science Publishers, Leiden, The Netherlands, 1993); van Gunsteren *et al.*, 1990 *Angew. Chem. Int. Ed. Engl.* 29: 992-1023; van Gunsteren *et al.*, 1992 *Eur. J. Biochem.* 204: 947-961; Torda and W.F. van Gunsteren. *Molecular Modeling Using Nuclear Magnetic Resonance Data*, IN REVIEWS IN COMPUTATIONAL CHEMISTRY,

Vol. III, 143-172 (K. B. Lipkowitz *et al.*, eds., VCH Publishers, Inc. New York, 1992); van Gunsteren, *Molecular dynamics and stochastic dynamics simulation: A primer*. in COMPUTER SIMULATION OF BIOMOLECULAR SYSTEMS, THEORETICAL AND EXPERIMENTAL APPLICATIONS 3-36; van Gunsteren *et al.*, *Computation of free energy in practice: choice of approximations and accuracy limiting factors*. IN COMPUTER SIMULATION OF BIOMOLECULAR SYSTEMS, THEORETICAL AND EXPERIMENTAL APPLICATIONS 315-348; Smith *et al.*, *Methods for the evaluation of long-range electrostatic forces in computer simulations of molecular systems*, IN COMPUTER

SIMULATION OF BIOMOLECULAR SYSTEMS, THEORETICAL AND EXPERIMENTAL APPLICATIONS 182-212; van Gunsteren *et al.*, *Accounting for Molecular Mobility in Structure Determination Based on Nuclear Magnetic Resonance Spectroscopic and X-Ray Diffraction Data*, in *Methods in ENZYMOLOGY: NUCLEAR MAGNETIC*
5 *RESONANCE*, Vol. 239 619-654 (T.L. James *et al.*, eds., Academic Press, New York, 1994); van Gunsteren *et al.*, 1994 *Quart. Rev. Biophysics* 27: 435-481; van Gunsteren *et al.*, 1995 *Biomolecular Modelling: Overview of Types of Methods to Search and Sample Conformational Space*, in *PROCEEDINGS OF THE 1ST EUROPEAN CONFERENCE ON COMPUTATIONAL CHEMISTRY*, AMERICAN INSTITUTE OF PHYSICS
10 CONF. PROC. 330: 253-268; and van Gunsteren *et al.*, 1995 *Computer Phys. Communications* 91: 305-319).

(K) The structure is then manually refined based on sequence alignment and structural elements using Deep View, the SWISS-PDB Viewer (Guex *et al.*, 1997 *Electrophoresis* 18: 2714-23).

15
Uses of structural model. Having a three-dimensional model of the protein domain allows one to appreciate the context of the HBM mutation, or other HBM-like mutations, in three dimensional space. This facilitates the prediction of possible mechanisms of action in a manner impossible in any other way because one cannot predict from primary sequence alone
20 the proximity of one distant (in sequence space) amino acid to another. This also allows one to make functional predictions that can then be tested using molecular/cellular biology techniques to further refine the model and validate the target. For example, the structural model of the G171V mutation predicted an alteration in side chain-side chain interactions, so we were able to use this information to successfully predict similar mutations in other amino
25 acids that would have analogous structural effects. Overlaying other functional data on this model would also able highlight other potentially important regions of the protein domain. Additionally, such models show the accessible residues, which would be useful for developing small compounds, antibodies and the like which recognize and bind to this region.

Results. Based on the space filling hypothesis set forth in Section 3 above, two residues from blade 6 of propeller 1 were identified as being in structurally equivalent locations to residue 171 in blade 4. Accordingly, substitutions with valine were engineered: F241V and A242V.

5 Residues that were predicted to be accessible to the interior surface of the propeller were identified in both blades 3 and 5 of Zmax1 (LRP5). Original predictions from the Springer model identified G199 on blade 5 and E128 on blade 3 as the equivalent position to residue 171 of blade 4 (HBM). Using the more sophisticated model described above, G199 and E128 are not predicted to be in the equivalent position to residue 171 of blade 4. Based
10 on the new, more sophisticated model, the following residues were chosen for valine substitutions:

<u>Blade No.</u>	<u>Mutations</u>
5	L200V, T201V and I202V
3	S127V

15

Another residue T125 was selected for a conservative substitution with T125S as well as a less conservative substitution of T125G.

20 The role of the YWTD repetitive motifs in the β -propeller was examined by carrying out an alanine replacement scan over this region in Blade 4. These repeats are predicted to create the β -sheet blades of the propeller. Based on this model data, the following additional mutations were prepared in propeller 1: Y164A, M165A, Y166A, W167A, T168A, and D169A. These substitutions would be predicted to be detrimental to the β -propeller structure.

25 The spacing filling or occupied space model was further examined by mutating a residue predicted by the above methods to reside on the exterior surface of propellar 1, blade 4. This mutation, K215V (K \rightarrow V at position 215), was selected because it was believed to not produce an HBM-like effect.

Based on this information, the following model of the HBM G171V mutation was obtained. Specifically, substitutions of valine at specific locations within propeller 1 could result in the creation of hydrophobic patches on the surface of the propeller. For example,

the V at residue 171 of the HBM mutant comes in close proximity to L150 on the adjacent strand, potentially creating a hydrophobic patch that was not present with the wild type G171. An example of this prediction, and others are presented in a separate file (Fig. 52).

Based on this model, a mutation of L150G which would remove the hydrophobic side chain at this site may not produce an HBM effect. In addition, a double mutant of G171V (HBM) and L150G may not produce an HBM effect due to the elimination of the hydrophobic side chain interactions. The model can also be used to test this prediction on a different propeller 1 blade. The substitution A214V was shown to result in an HBM effect. I194 is the residue that is in the equivalent position to A214 on blade 5 as L150 is to G171 on blade 4. While the mutation of I194G alone may not be predicted to result an HBM effect, the double mutant of I194G and A214V may no longer permit a hydrophobic patch to form and as a result not generate the HBM effect observed with A214V. Additional analyses of the residues in proximity to A214 that may contribute to the formation of hydrophobic patches identified P240 and F241. Mutations of either P240G or F241G together with A214V may no longer permit the formation of hydrophobic patches and may eliminate the HBM effect observed with A214V alone.

Lastly, the new model was used to investigate specific aspects of the other three propellers. The relevance of residues R494, R570 and V667 in the three other propellers was investigated. These three residues were chosen since they have been reported to be the site of osteoporosis pseudoglioma disease causing missense mutations: R494Q, R570W and V667M (Gong *et al.*, 2001 *Cell* 107: 513-523). Based on the structural location of these residues on the outer surface of the propeller, the model would not predict that these substitutions are either gain- or loss-of-function substitutions (Fig. 53A-C). For example, the mutations related to osteoporosis-pseudoglioma syndrome (OPPG) result in patients having very low bone mass and who are prone to developing fractures and deformation. This genetic condition is known now to result from mutations in LRP5 (Gong *et al.*, 2001 *Cell* 107: 513-23). Based on this information, this mutation would be categorized as a loss in function mutation.

As a result, we have engineered these receptor variant to functionally test. These new

receptor variants were generated exactly as described above with the following oligonucleotides:

5	L150GF	GAAGGTGCTCTTCTGGCAGGACGGTGACCAGCCGAGGGCC
	L150GR	GGCCCTCGGCTGGTCACCGTCCTGCCAGAAGAGCACCTTC
10	I194GF	GATCATTGTGGACTCGGACGGTTACTGGCCCAATGGACTG
	I194GR	CAGTCCATTGGGCCAGTAACCGTCCGAGTCCACAATGATC
15	L200VF	CATTTACTGGCCCAATGGAGTGACCATCGACCTGGAGGAGCAGAAGC
	L200VR	GCTTCTGCTCCTCCAGGTCGATGGTCACTCCATTGGGCCAGTAAATG
20	T201VF	CATTTACTGGCCCAATGGACTGGTCATCGACCTGGAGGAGCAGAAGC
	T201VR	GCTTCTGCTCCTCCAGGTCGATGACCAGTCCATTGGGCCAGTAAATG
25	I202VF	CATTTACTGGCCCAATGGACTGACCGTCGACCTGGAGGAGCAGAAGC
	I202VR	GCTTCTGCTCCTCCAGGTCGACGGTCAGTCCATTGGGCCAGTAAATG
30	T125SF	GGCAAGAAGCTGTACTGGTCGGACTCAGAGACCAACCGCATC
	T125SR	GATGCGGTTGGTCTCTGAGTCCGACCAGTACAGCTTCTTGCC
35	T125GF	GGCAAGAAGCTGTACTGGGGGGACTCAGAGACCAACCGCATC
	T125GR	GATGCGGTTGGTCTCTGAGTCCCCCAGTACAGCTTCTTGCC
40	S127VF	GAAGCTGTACTGGACGGACGTAGAGACCAACCGCATCGAGGTG
	S127VR	CACCTCGATGCGGTTGGTCTCTACGTCCGTCCAGTACAGCTTC
45	F241VF	CCTGACGCACCCCGTCGCCCTGACGCTCTCCGGGGACACTC
	F241VR	GAGTGTCCCCGGAGAGCGTCAGGGCGACGGGGTGCGTCAGG
50	A242VF	CCTGACGCACCCCTTCGTCTGACGCTCTCCGGGGACACTC
	A242VR	GAGTGTCCCCGGAGAGCGTCAGGACGAAGGGGTGCGTCAGG
55	K215VF	CTACTGGGCTGACGCCGTGCTCAGCTTCATCCACCGTGCC
	K215VR	GGCACGGTGGATGAAGCTGAGCACGGCGTCAGCCCAGTAG
60	Y164AF	CTTGACCCCGCTCACGGGGCCATGTACTGGACAGACTGG
	Y164AR	CCAGTCTGTCCAGTACATGGCCCCGTGAGCGGGGTCCAAG
65	M165AF	GGACCCCGCTCACGGGTACGCGTACTGGACAGACTGGGGTG
	M165AR	CACCCAGTCTGTCCAGTACGCGTACCCGTGAGCGGGGTCC
70	Y166AF	CCCGCTCACGGGTACATGGCCTGGACAGACTGGGGTGAGAC
	Y166AR	GTCTACCCCAAGTCTGTCCAGGCCATGTACCCGTGAGCGGG
75	W167AF	CGCTCACGGGTACATGTACGCGACAGACTGGGGTGAGACGC
	W167AR	GCGTCTACCCCAAGTCTGTGCGGTACATGTACCCGTGAGCG

	T168AF	GCTCACGGGTACATGTACTGGGCAGACTGGGGTGAGACGCC
	T168AR	GGCGTCTCACCCCAGTCTGCCCAGTACATGTACCCGTGAGC
5	D169AF	CACGGGTACATGTACTGGACAGCCTGGGGTGAGACGCCCCG
	D169AR	CGGGGCGTCTCACCCCAGGCTGTCCAGTACATGTACCCGTG
	R494QF	CAACTTGGATGGGCAGGAGCAGCGTGTGCTGGTCAATGCCTC
	R490QR	GAGGCATTGACCAGCACACGCTGCTCCTGCCCATCCAAGTTG
10	R570WF	GCAGCGCCGCAGCATCGAGTGGGTGCACAAGGTCAAGGCCAG
	R570WR	CTGGCCTTGACCTTGTGCACCCACTCGATGCTGCGGCGCTGC
	V667MF	CGAGACCAATAACAACGACATGGCCATCCCGCTCACGGGCG
15	V667MR	CGCCCGTGAGCGGGATGGCCATGTCGTTGTTATTGGTCTCG

Point mutations in other molecules with beta-propellers have been described in the literature. For example, alpha4-integrin (Guerrero-Esteo *et al.*, 1998 *FEBS Letter* 429: 123-8) has had two mutations introduced, which resulted in altered protein function (*i.e.*, G130R and G190S). Both of these glycines are in structurally equivalent locations to G171 of LRP5, just located on different blades of the beta-propeller at the upper surface. Using the model discussed above, we mapped these residues of the alpha4-integrin and confirmed their placement. The result of these substitutions in the alpha4 integrin was loss of ligand binding and reduced affinity to its heterodimer binding partner: integrin-beta1, such that the alpha4-beta1 ($\alpha 4 \beta 1$) heterodimer cannot be formed. Examples of modulation of other residues in integrin-alpha4 with functional consequences are referenced within Guerrero-Esteo *et al.* (1998). Thus, these data support the notion that disruption of beta-propeller structure can result in significant and dramatic effects on receptor function.

All references cited are herein incorporated by reference in their entirety for all purposes. The following applications are also incorporated by reference in their entirety herein: U.S. Application Nos. 09/543,771 and 09/544,398 filed on April 5, 2000, which are a continuation-in-part of Application No. 09/229,319, filed January 13, 1999, which claims benefit of U.S. Provisional Application No. 60/071,449, filed January 13, 1998, and U.S. Provisional Application No. 60/105,511, filed October 23, 1998. Additionally This application claims priority of Application Nos. 60/290,071 filed May 11, 2001; 60/291,311

filed May 17, 2001; 60/353,058 filed February 1, 2002, and 60/361,293 filed March 4, 2002
the texts of which are herein incorporated by reference in their entirety for all purposes.

CLAIMS

What is claimed is:

1. A nucleic acid comprising a mutation in LRP5 or LRP6 which results in a HBM-like phenotype when expressed in a cell, wherein said HBM-like phenotype results in bone mass modulation and/or lipid level modulation.
2. A LRP5 nucleic acid comprising at least one mutation of Tables 2 or 3 or which results in a mutation of one of the following G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V when expressed in a cell, and wherein expression of the nucleic acid in a subject results in bone mass modulation and/or lipid level modulation.
3. A LRP6 nucleic acid comprising at least one of Tables 2 or 3 or which results in a mutation of one of the following G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V when expressed in a cell, and wherein said mutation occurs in LRP6 in a position equivalent to LRP5 such that expression of the nucleic acid in a subject results in bone mass modulation and/or lipid level modulation.
4. A LRP5 nucleic acid encoding at least one amino acid mutation in propeller 1.
5. The LRP5 nucleic acid of claim 4, wherein said nucleic acid encodes for at least one mutation selected from the group consisting of: G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, and S127V.
6. A LRP6 nucleic acid encoding at least one amino acid mutation in propeller 1.
7. The LRP6 nucleic acid of claim 6, wherein said nucleic acid encodes for at least one mutation in an equivalent position of LRP6 selected from the group of LRP5 mutations consisting of: G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, and S127V.
8. The nucleic acid of claims 1-7, wherein the mutation is G171V, A214V,

A65V, M282V, G171K, G171F, G171I or G171Q in LRP5 or in an equivalent domain in LRP6.

9. A polypeptide encoded by a nucleic acid of claims 1 to 8, wherein said polypeptide when expressed in a cell modulates Wnt signaling, Dkk activity, LRP5 activity and/or LRP6 activity.

10. The polypeptide of claim 9, wherein said polypeptide modulates bone mass and/or lipid levels when expressed in a subject.

11. A vector comprising a nucleic acid of claims 1 to 8.

12. A cell comprising the vector of claim 11.

13. The cell of claim 12, wherein said cell is a cancer cell, a liver cell or a bone cell.

14. A polypeptide or biologically active fragment thereof derived from LRP5 comprising at least one amino acid change of Table 2, G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V.

15. The polypeptide or biologically active fragment of claim 14, wherein said protein or biologically active fragment comprises at least one amino acid change of G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V.

16. A polypeptide or biologically active fragment thereof derived from LRP6 comprising at least one amino acid change of Table 2, G171V, A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V when the amino acid change is expressed in an equivalent position in LRP6.

17. The polypeptide or biologically active fragment of claim 16, wherein said protein or biologically active fragment comprises at least one amino acid change of G171V,

A214V, A65V, M282V, G171K, G171F, G171I, G171Q, L200V, T201V, I202V, or S127V.

18. A polypeptide or biologically active fragment which when expressed in a subject has a HBM-like phenotype comprising at least one amino acid mutation in propeller 1 with the proviso that if the mutation is GV171 that a second mutation which maintains the HBM-like phenotype is also present.

19. The polypeptide or biologically active fragment of claim 18, wherein said polypeptide is derived from LRP5, LRP6 or HBM.

20. The polypeptide or biologically active fragment of claims 14-19, wherein the amino acid change is G171V, A214V, A65V, M282V, G171K, G171F, G171I or G171Q of LRP5 or an equivalent position in LRP6.

21. An antibody or immunogenic fragment thereof which binds to a polypeptide or antigenic fragment thereof of claims 14 to 20.

22. The antibody or immunogenic fragment thereof of claim 21, wherein the antibody is a monoclonal antibody, a chimeric antibody, a bispecific antibody, a humanized antibody, a primatized® antibody, a human antibody, or a labeled antibody.

23. An antibody which binds to a polypeptide comprising
²⁰⁸KLYWADAKLSFIHRAN²²³, ²⁷⁷ALYSPMDIQVLSQER²⁹¹, ⁶¹GLEDAAAVDFQFSKGA⁷³,
²³⁴EGSLTHPFALTLSG²⁴⁷, ²⁴⁹TLYWTDWQTRSIHACN²⁶⁴, ¹⁴⁴VLFWQDLDQPRAI¹⁵⁶,
¹⁹⁴IYWPNGLTIDLEEQKLY²¹⁰, ³⁴LLLFANRRDVRLVD⁴⁷,
⁷⁵GAVYWTDVSEEAIKQ⁸⁹, ¹²¹KLYWTDSETNRIEVA¹³⁵ of LRP5 or an equivalent domain on LRP6 or variants thereof.

24. An antibody which binds to a polypeptide comprising
⁹⁶⁹LILPLHGLRNVKAIDYDPLDKFIYW⁹⁹³, ⁹⁸⁹KFIYWVDGRQNIKRAKDDGTQPFVL¹⁰¹³,
¹⁰⁰⁹QPFVLTSLSQGQNPDRQPHDLSIDI¹⁰³³, ¹⁰²⁹LSIDIYSRTLFWTCEATNTINVHRL¹⁰⁵³,
¹⁰⁴⁹NVHRLSGEAMGVVLRGDRDKPRAIV¹⁰⁷³, ¹²⁵³CGEPPTCSPDQFAC¹²⁶⁶,

¹²⁷⁸WRCDGFPECDDQSDEEGC¹²⁹⁵, ¹³¹⁶RCDGEADCQDRSDEADC¹³³²,
¹³⁷⁰CEITKPPSDDSPA¹³⁸³ of LRP5 or an equivalent domain on LRP6 or variants thereof,
 wherein said antibody modulates binding of LRP5, HBM, LRP6 or variants thereof to Dkk.

25. A method of diagnosing a HBM-like phenotype in a subject comprising
 (A) obtaining a biological sample from the subject;
 (B) exposing the sample to an antibody or an immunogenic fragment of claims 21-24; and
 (C) detecting whether the antibody bound a protein from the biological sample from the subject to determine whether the subject has a HBM-like phenotype.

26. A composition for modulating bone mass and/or lipid levels in a subject comprising a therapeutically effective amount of an antibody of claims 21-24 and a pharmaceutically acceptable carrier.

27. An antibody or an immunogenic fragment thereof of claims 21-24, wherein the antibody or immunogenic fragment thereof can (1) discriminate between LRP5 and an HBM-like protein, (2) discriminate between LRP6 and an HBM-like protein, or (3) discriminate between HBM and an HBM-like protein.

28. A transgenic animal having somatic and/or germ cells comprising a nucleic acid which comprises a promoter region that directs protein expression in animal and/or human cells operably linked to a nucleic acid of claims 1 to 8, and wherein said transgenic animal has at least three bone parameters modulated by the expression of said nucleic acid.

29. The transgenic animal of claim 28, wherein the promoter region is selected from the group consisting of CMV, RSV, SV40, and EF-1a, CMV β Actin, histone, type I collagen, TGF β 1, SX2, cfos/cjun, Cbfa1, Fra/Jun, Dlx5, osteocalcin, osteopontin, bone sialoprotein, and collagenase promoter regions.

30. A transgenic animal having somatic and/or germ cells comprising a nucleic

acid which comprises a sequence which encodes a polypeptide of claims 9 or 10, and wherein the nucleic acid further comprises an operably linked promoter region that directs protein expression in animal and/or human cells, and wherein the transgenic animal has at least three bone parameters modulated by the expression of said nucleic acid.

31. An animal embryo comprising a nucleic acid comprising a promoter region that directs protein expression in animal and/or human cells operably linked to a nucleic acid of claims 1 to 8.

32. An animal or human cell transfected with a nucleic acid which comprises a promoter region that directs protein expression in animal and/or human cells operably linked to a nucleic acid of claims 1 to 8.

33. The transgenic animal of claim 28, wherein a human HBM-like protein is expressed and passes the human *HBM*-like gene to its offspring.

34. A transgenic animal produced from the transgenic animal of claim 33 or its offspring.

35. An animal model for the study of bone density modulation and/or lipid level modulation comprising a first group of animals composed of the transgenic animal of claim 28 and a second group of control animals.

36. The animal model of claim 35, wherein the group of transgenic animals have cells which comprise a nucleic acid encoding human an HBM-like protein or biologically active polypeptide fragment thereof.

37. An animal model for the study of bone density modulation comprising a first group of animals composed of the transgenic animals of claim 30 and a second group of control animals.

38. A method of identifying agents which modulate the activity of an HBM-like

nucleic acid comprising:

- (a) transfecting a cell with a vector of claim 11;
- (b) exposing the transfected cell of step (a) to a compound; and
- (c) determining whether the compound modulates the activity of the HBM-like nucleic acid.

39. A method of identifying agents which modulate the activity of an HBM-like protein comprising:

- (a) transfecting a cell with a vector of claim 11;
- (b) exposing the transfected cell of step (a) to a compound; and
- (c) determining whether the compound modulates the activity of the HBM-like protein.

40. The method of claim 39, wherein the compound is a hormone, a growth factor, a peptide, RNA, shRNA, siRNA, DNA, a mineral, a vitamin, a natural product, or a synthetic organic compound.

41. A method for identifying compounds which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

- (a) transfecting cells with constructs containing a nucleic acid of claims 1 to claim 8;
- (b) assessing changes in expression of a reporter element linked to a Wnt-responsive promoter; and
- (c) identifying as a Dkk/Wnt interaction modulating compound any compound which alters reporter gene expression compared with cells transfected with a Dkk construct alone.

42. The method according to Claim 41, wherein the cells are cancer cells, liver cells or bone cells.

43. The method according to Claim 42, wherein the cells are U2-OS, HOB-03-CE6, or HEK293 cells.

44. The method according to Claim 41, wherein the reporter element used is TCF-luciferase, tk-Renilla, or a combination thereof.

45. A method of diagnosing a subject as expressing a nucleic acid comprising a nucleotide change of Tables 2 or 3, the method comprising the steps of:

- (A) obtaining a biological sample from the subject; and
- (B) assaying for the presence of the nucleotide change which results in HBM phenotype.

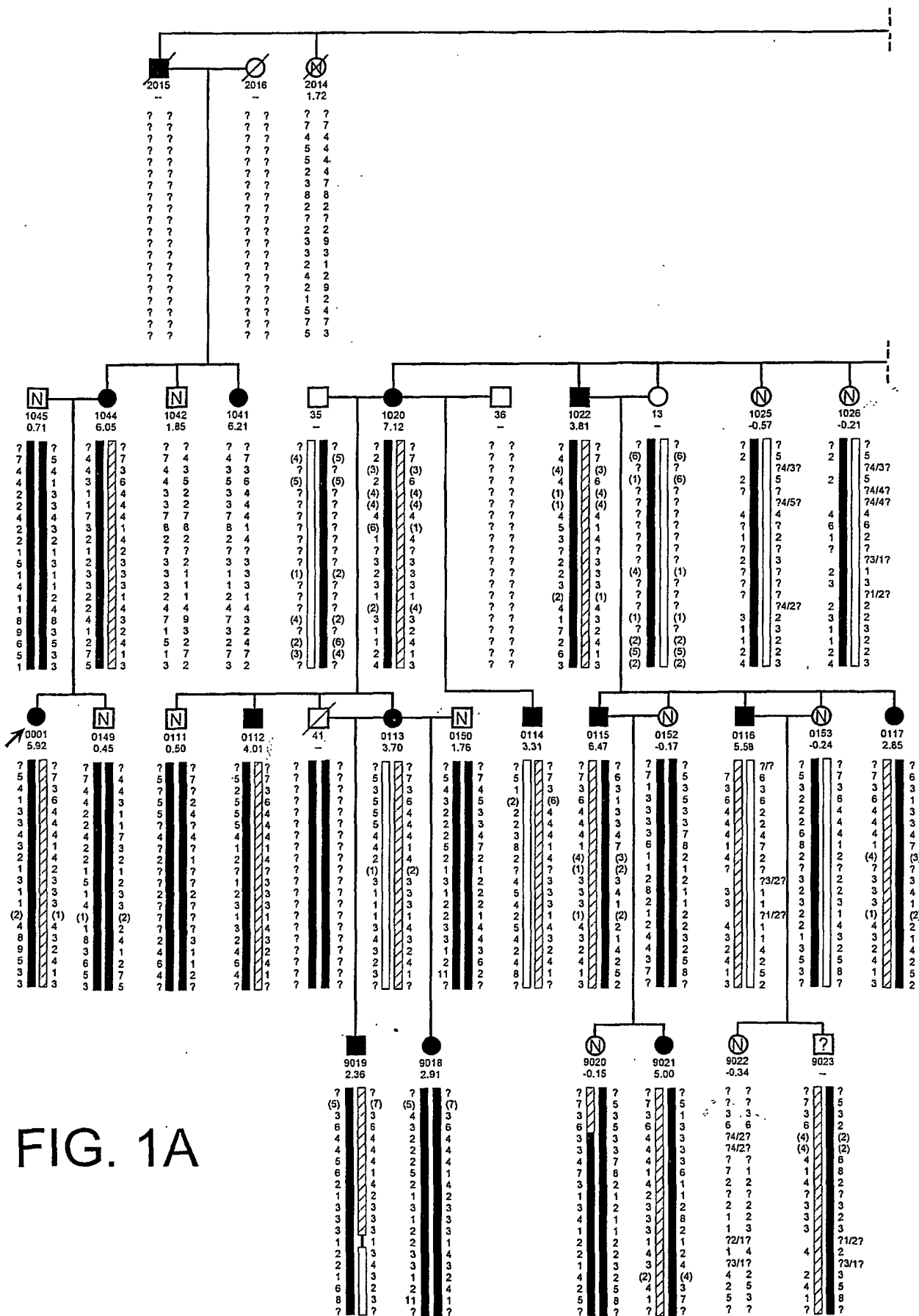
46. A method for identifying agents which modulate LRP5, LRP6, or HBM comprising the steps of:

- (A) providing cells according to claim 32;
- (B) exposing the cells to a test compound; and
- (C) measuring the expression of LRP5, LRP6 or HBM respectively.

47. The method for identifying agents of claim 42, further comprising a step of determining whether the agent further modulates bone mass and/or lipid levels.

48. An agent identified by the method of claims 46 or 47.

1/89



2/89

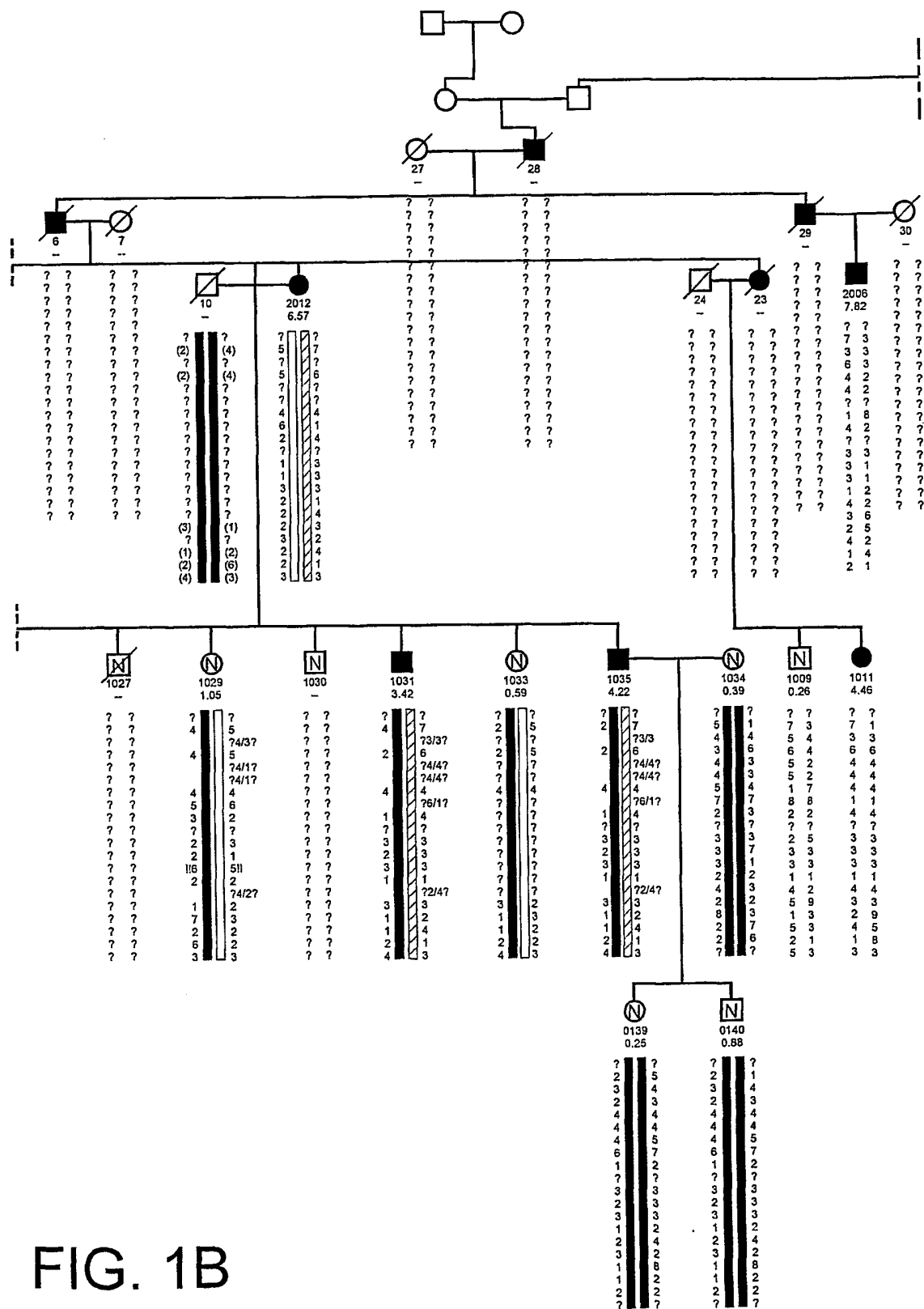


FIG. 1B

FIG. 2A
BAC/STS Map of the HBM Region

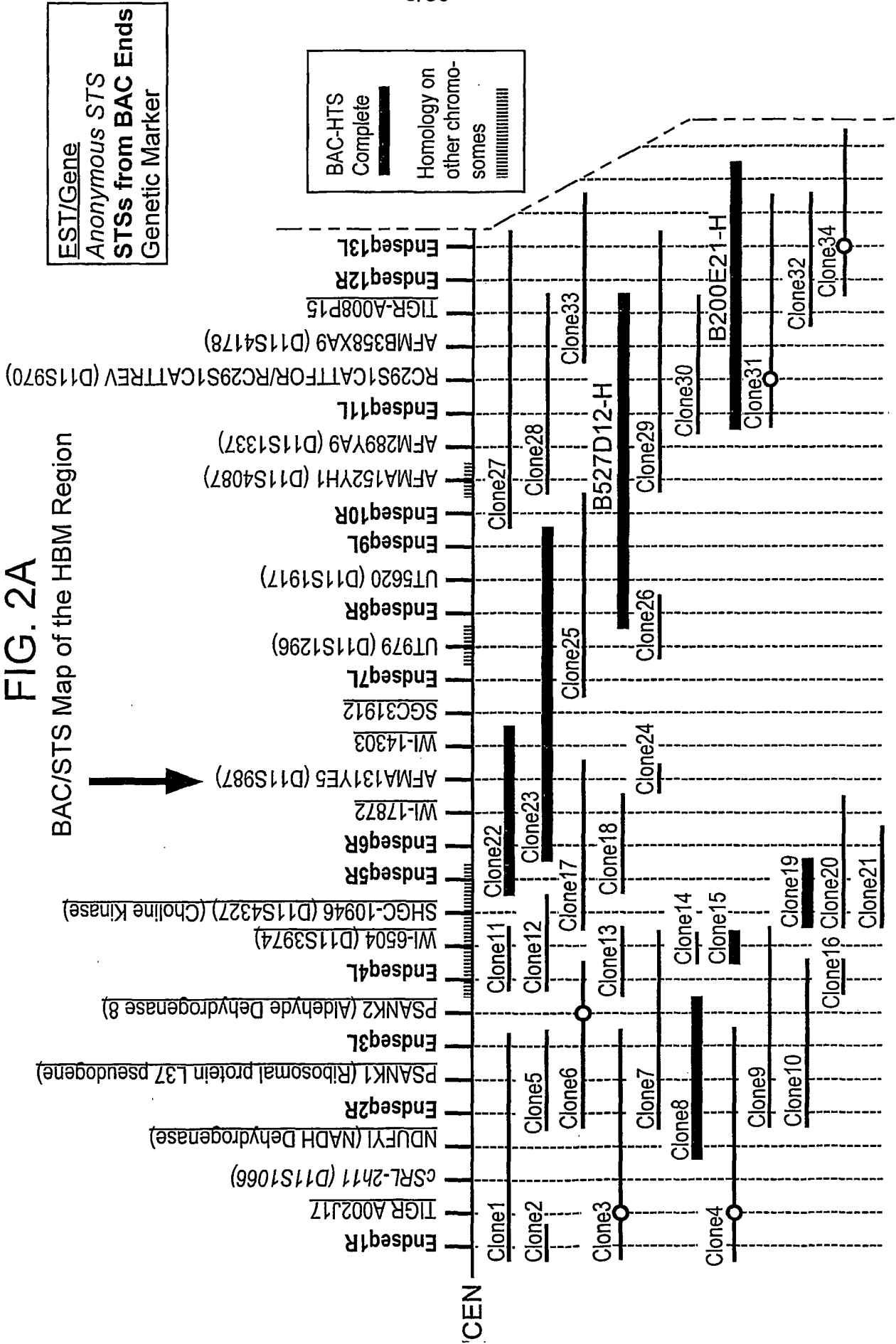
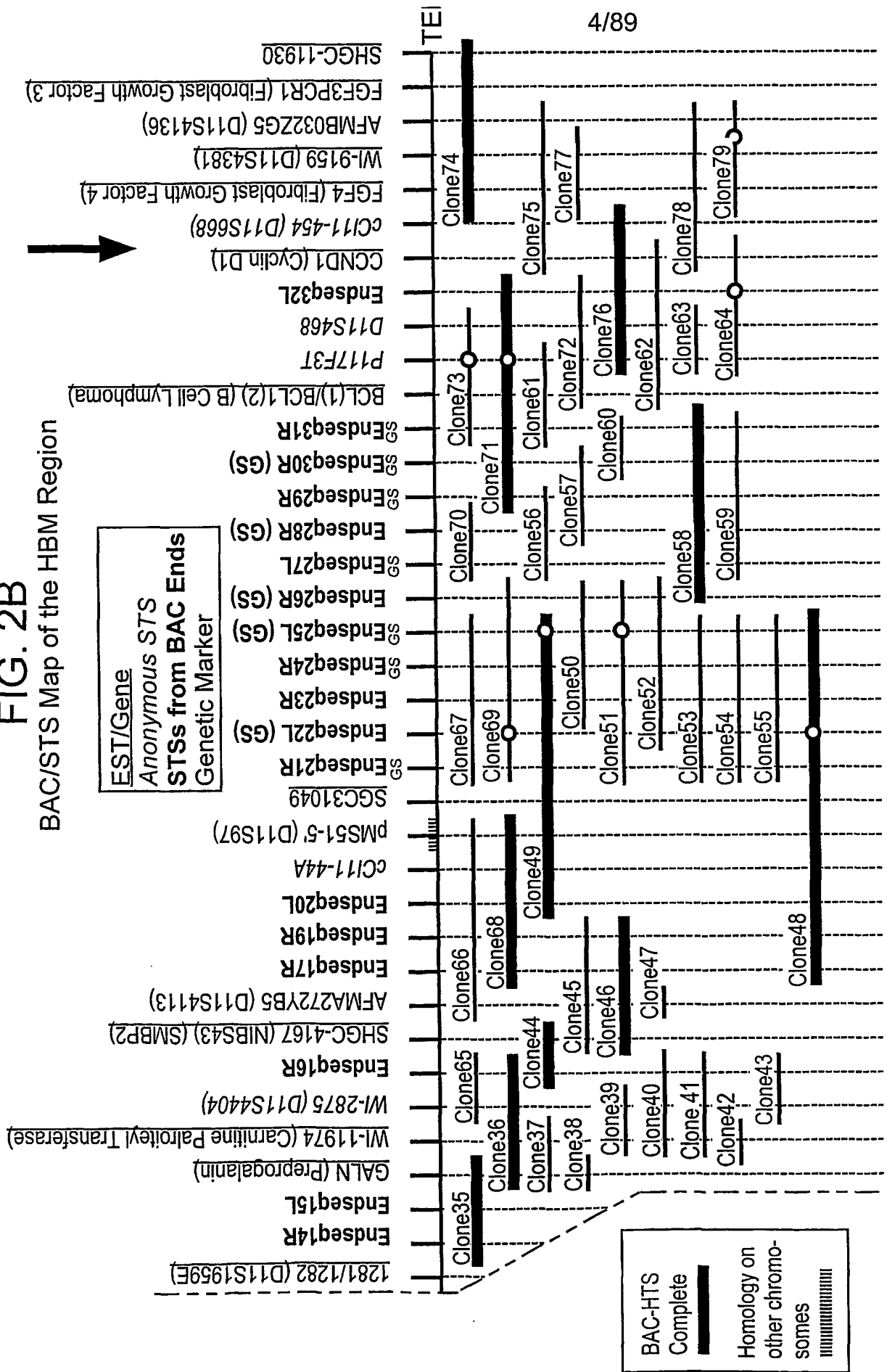


FIG. 2B
BAC/STS Map of the HBM Region



5/89

Exon 1

ACTAAAGCGCCGCCGCCGCCATGGAGCCCGAGTGAGCGCGGGCGCG
GGCCCGTCCGGCCGCCGGACAACATGGAGGCAGCGCCGCCCGGGCCG
CCGTGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGGCGCTGTGCGGC
TGCCCGGCCCGGCCGCCGCC

Exon 2 Coordinates: 527d12_Contig308G 30944-30549

gccccacagCCTCGCCGCTCCTGCTATTTGCCAACC GCCGGGACGTACGGC
TGGTGGACGCCGGCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAGC
GGCCTGGAGGATGCGGCCGCACTGGACTTCCAGTTTTCCAAGGGAGC
CGTGTACTGGACAGACGTGAGCGAGGAGGCCATCAAGCAGACCTACCT
GAACCAGACGGGGGCCGCCGTGCAGAACGTGGTCATCTCCGGCCTGG
TCTCTCCCGACGGCCTCGCCTGCGACTGGGTGGGCAAGAAGCTGTACT
GGACGGACTCAGAGACCAACCGCATCGAGGTGGCCAACCTCAATGGC
ACATCCCGGAAGGTGCTCTTCTGGCAGGACCTTGACCAGCCGAGGGCC
ATCGCCTTGGACCCCGCTCACGGgtaaaccctgctg

... 9408 nt ...

Exon 3 Coordinates: 527d12_Contig308G 21141-20945

ccccgtcacagGTACATGTACTGGACAGACTGGGGTGAGACGCCCCGGATTG
AGCGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTGTGGACTCG
GACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAAG
CTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCAACCTG
GACGGCTCGTTCCGgtaggtaccac

... 6094 nt ...

Exon 4 Coordinates: 527d12_Contig308G 15047-14850

tccctgactgcagGCAGAAGGTGGTGGAGGGCAGCCTGACGCACCCCTTCGCC
CTGACGCTCTCCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGC
TCCATCCATGCCTGCAACAAGCGCACTGGGGGGAAGAGGAAGGAGAT
CCTGAGTGCCCTATACTACCCATGGACATCCAGGTGCTGAGCCAGGA
GCGGCAGCCTTTCTgtgagtgccg

... 1827 nt ...

Exon 5 Coordinates: 527d12_Contig308G 13220-13088

tttctcagTCCACACTCGCTGTGAGGAGGACAATGGCGGCTGCTCCCACCTG
TGCTGTGTCCCCAAGCGAGCCTTTCTACACATGCGCCTGCCCCACG
GGTGTGCAGCTGCAGGACAACGGCAGGACGTGTAAGGCAGgtgaggcggtgg
gacg

FIG. 3A

6/89

... 20923 nt ...

Exon 6 Coordinates: 527d12_Contig309G 7705-8100

ctccacagGAGCCGAGGAGGTGCTGCTGCTGGCCCCGGCGGACGGACCTAC
GGAGGATCTCGCTGGACACGCCGGA CTTCACCGACATCGTGCTGCAGG
TGGACGACATCCGGCACGCCATTGCCATCGACTACGACCCGCTAGAGG
GCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATCCGCAGGGCG
TACCTGGACGGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAA
CGACCCCGATGGCATCGCGGTCTGACTGGGTGGCCCCGAAACCTCTACTG
GACCGACACGGGCACGGACCGCATCGAGGTGACGCGCCTCAACGGCA
CCTCCCGCAAGATCCTGGTGTCTGGAGGACCTGGACGAGCCCCGAGCC
ATCGCACTGCACCCCGTGATGGGgtaagacgggc

..... 3211 nt

Exon 7 Coordinates: 527d12_Contig309G 11311-11482

ttcttctccagCCTCATGTACTGGACAGACTGGGGAGAGAACCCTAAAATCGA
GTGTGCCAACTTGGATGGGCAGGAGCGGCGTGTGCTGGTCAATGCCTC
CCTCGGGTGGCCCAACGGCCTGGCCCTGGACCTGCAGGAGGGGGAAGC
TCTACTGGGGAGACGCCAAGACAGACAAGATCGAGgtgaggctcctgtgg

..... 13445 nt

Exon 8 Coordinates: 527d12_Contig309G 24927-25143

ccgtctgcagGTGATCAATGTTGATGGGACGAAGAGGGCGGACCCTCCTGGA
GGACAAGCTCCCGCACATTTTCGGGTTACGCTGCTGGGGGACTTCAT
CTACTGGACTGACTGGCAGCGCCGCAGCATCGAGCGGGTGCACAAGG
TCAAGGCCAGCCGGGACGTCATCATTGACCAGCTGCCCCGACCTGATGG
GGCTCAAAGCTGTGAATGTGGCCAAGGTCGTCGgtgagtcgggggggtc

....2826 nt

Exon 9 Coordinates: 527d12_Contig309G 27969-28256

gttcgcttcagGAACCAACCCGTGTGCGGACAGGAACGGGGGGGTGCAGCCA
CCTGTGCTTCTTCACACCCACGCAACCCGGTGTGGCTGCCCCATCGG
CCTGGAGCTGCTGAGTGACATGAAGACCTGCATCGTGCTGAGGCCTT
CTTGGTCTTCACCAGCAGAGCCGCCATCCACAGGATCTCCCTCGAGAC
CAATAACAACGACGTGGCCATCCCGCTCACGGGCGTCAAGGAGGCCTC
AGCCCTGGACTTTGATGTGTCCAACAACCACATCTACTGGACAGACGT
CAGCCTGAAGgtagcgtgggc

.....3102.....

FIG. 3B

7/89

Exon 10 Coordinates: 527d12_Contig309G 31358-31582

cctgctgccagACCATCAGCCGCGCCTTCATGAACGGGAGCTCGGTGGAGCA
CGTGGTGGAGTTTGGCCTTGACTACCCCGAGGGCATGGCCGTTGACTG
GATGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAATCGA
AGTGGCGCGGCTGGACGGGCAGTTCCGGCAAGTCCTCGTGTGGAGGG
ACTTGGACAACCCGAGGTCGCTGGCCCTGGATCCCACCAAGGGgtaagtgtt
tgcctgtc

.....1297 nt.....

Exon 11 Coordinates: 527d12_Contig309G 32879-33064

gtgccttcagCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCG
TGCGGGCCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAG
GTGGGCCCGGGCCAACGACCTCACCATTGACTACGCTGACCAGCGCCTC
TACTGGACCGACCTGGACACCAACATGATCGAGTCGTCCAACATGCTG
Ggtgagggccgggct

.....2069 nt.....

Exon 12 Coordinates: 527d12_Contig309G 35133-35454

gtgtcatgcagGTCAGGAGCGGGTCGTGATTGCCGACGATCTCCCGCACCCG
TTCGGTCTGACGCAGTACAGCGATTATCTACTGGACAGACTGGAAT
CTGCACAGCATTGAGCGGGCCGACAAGACTAGCGGCCGGAACCGCAC
CCTCATCCAGGGCCACCTGGACTTCGTGATGGACATCCTGGTGTTC
CTCCTCCCGCCAGGATGGCCTCAATGACTGTATGCACAACAACGGGCA
GTGTGGGCAGCTGTGCCTTGCCATCCCCGGCGGCCACCGCTGCGGCT
GCGCCTCACACTACACCCTGGACCCAGCAGCCGCAACTGCAGCCgtaag
tgcctcatggt

.....2006 nt.....

Exon 13 Coordinates: 527d12_Contig309G 37460-37659

gcctcctctaCGCCCACCACCTTCTTGCTGTTTCAGCCAGAAATCTGCCATCAG
TCGGATGATCCCGGACGACCAGCACAGCCCGGATCTCATCCTGCCCCCT
GCATGGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAA
GTTTCATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAAGGA
CGACGGGACCCAGgcaggtgccctgtgg

.....6965 nt.....

FIG. 3C

8/89

Exon 14 Coordinates: 527d12_Contig309G 44624-44832

ctttgtcttacagCCCTTTGTTTTGACCTCTCTGAGCCAAGGCCAAAACCCAGAC
AGGCAGCCCCACGACCTCAGCATCGACATCTACAGCCGGACACTGTTC
TGGACGTGCGAGGCCACCAATACCATCAACGTCCACAGGCTGAGCGG
GGAAGCCATGGGGGTGGTGCTGCGTGGGGACCGCGACAAGCCCAGGG
CCATCGTCGTCAACGCGGAGCGAGGgtaggaggccaac

.....1404 nt.....

Exon 15 Coordinates: 527d12_Contig309G 46236-46427

ccaccctccgcagGTACCTGTACTTCACCAACATGCAGGACCGGGCAGCCAA
GATCGAACGCGCAGCCCTGGACGGCACCAGCGCGAGGTCCTCTTCA
CCACCGGCCTCATCCGCCCTGTGGCCCTGGTGGTGGACAACACACTGG
GCAAGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTGAGAGCTGT
GACCTGTCAGgtacgcgccccgg

.....686 nt.....

Exon 16 Coordinates: 527d12_Contig309G 47113-47322

ggctgcttcagGGGCCAACCGCCTGACCCTGGAGGACGCCAACATCGTGCA
GCCTCTGGGCCTGACCATCCTTGGCAAGCATCTCTACTGGATCGACCG
CCAGCAGCAGATGATCGAGCGTGTGGAGAAGACCACCGGGGACAAGC
GGACTCGCATCCAGGGCCGTGTGCCCCACCTCACTGGCATCCATGCAG
TGGAGGAAGTCAGCCTGGAGGAGTTCTgtacgtgggggc

.....3884 nt.....

Exon 17 Coordinates: 527d12_Contig309G 51206-51331

ttgtctttgcagCAGCCCACCCATGTGCCCCGTGACAATGGTGGCTGCTCCCACA
TCTGTATTGCCAAGGGTGATGGGACACCACGGTGCTCATGCCCAGTCC
ACCTCGTGCTCCTGCAGAACCTGCTGACCTGTGGAGgtaggtgtgacctaggtgc

....3905 nt.....

Exon 18 Coordinates: 527d12_Contig309G 55236-55472

gttctcctctgtccctccccagAGCCGCCACCTGCTCCCCGGACCAGTTTGCATGT
GCCACAGGGGAGATCGACTGTATCCCCGGGGCCTGGCGCTGTGACGG
CTTTCCCGAGTGCGATGACCAGAGCGACGAGGAGGGGCTGCCCCGTGT
GCTCCGCCGCCAGTTCCCCTGCGCGCGGGGTCAGTGTGTGGACCTGC
GCCTGCGCTGCGACGGCGAGGCAGACTGTCAGGACCGCTCAGACGAG
GTGGACTGTGACGgtgaggccctcc

.....3052 nt.....

FIG. 3D

9/89

Exon 19 Coordinates: 527d12_Contig309G 58524-58634

tctccttgagCCATCTGCCTGCCCAACCAGTTCCGGTGTGCGAGCGGCCAGT
GTGTCCTCATCAAACAGCAGTGCGACTCCTTCCCCGACTGTATCGACG
GCTCCGACGAGCTCATGTGTGgtgagccagctt

.....1448 nt.....

Exon 20 Coordinates: 527d12_Contig309G 60082-60319

gtttgtctctggcagAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACACA
GCAGTGCCATCGGGCCCGTCATTGGCATCATCCTCTCTCTCTTCGTCAT
GGGTGGTGTCTATTTTGTGTGCCAGCGCGTGGTGTGCCAGCGCTATGC
GGGGGCCAACGGGGCCCTTCCCGCACGAGTATGTCAGCGGGACCCCGC
ACGTGCCCCCTCAATTTTCATAGCCCCGGGCGGTTCCCAGCATGGCCCCCT
TCACAGgtaaggagcctgagatatggaa

....1095 nt.....

Exon 21 Coordinates: 527d12_Contig309G 61414-61552

cttccttgccagGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG
ATGGGGGGCCGGGGCGGGGTGCCCTCTACGACCGGAACACGTCAC
AGGGGCCTCGTCCAGCAGCTCGTCCAGCACGAAGGCCACGCTGTACCC
GCCGgtgaggggcggg

.....6513 nt.....

Exon 22 Coordinates: 527d12_Contig309G 68065-68162

ttggctctcctcagATCCTGAACCCGCCGCCCTCCCCGGCCACGGACCCCTCCC
TGTACAACATGGACATGTTCTACTCTTCAAACATTCCGGCCACTGCGA
GACCGTACAGgtaggacatcccctgcag

.....2273 nt.....

FIG. 3E

10/89

Exon 23 Coordinates: 527d12_Contig309G 70435-70901

tcaaacattccggccactgcgagaccgtacagGCCCTACATCATTCGAGGAATGGCGCCCC
CGACGACGCCCTGCAGCACCGACGTGTGTGACAGCGACTACAGCGCC
AGCCGCTGGAAGGCCAGCAAGTACTACCTGGATTTGAACTCGGACTCA
GACCCCTATCCACCCCCACCCACGCCCCACAGCCAGTACCTGTCGGCG
GAGGACAGCTGCCCCGCCCTCGCCCCGCCACCGAGAGGAGCTACTTCCAT
CTCTTCCC GCCCCTCCGTCCCCCTGCACGGACTCATCCTGACCTCGGC
CGGGCCACTCTGGCTTCTCTGTGCCCTGTAAATAGTTTTAAATATGAACAA
AGAAAAAATATATTTTATGATTAAAAAATAAATATAATTGGGATTTTAA
AAACATGAGAAATGTGAACTGTGATGGGGTGGGCAGGGCTGGGAGAACTT
TGTACAGTGGAGAAATATTTATAAACTTAATTTTGTAACA

FIG. 3F

Model for a LDL Receptor-Related protein, Zmax1

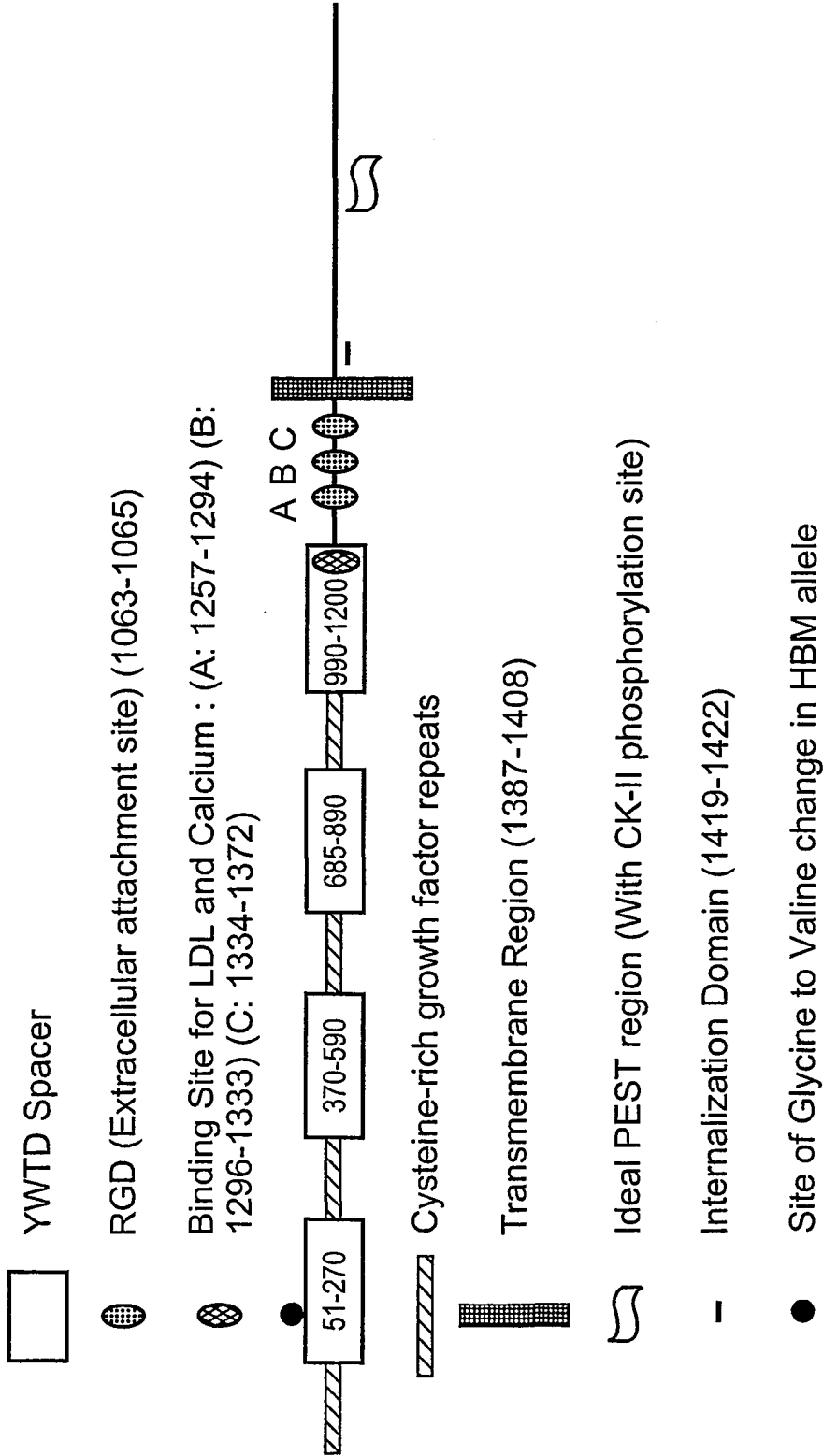


FIG. 4

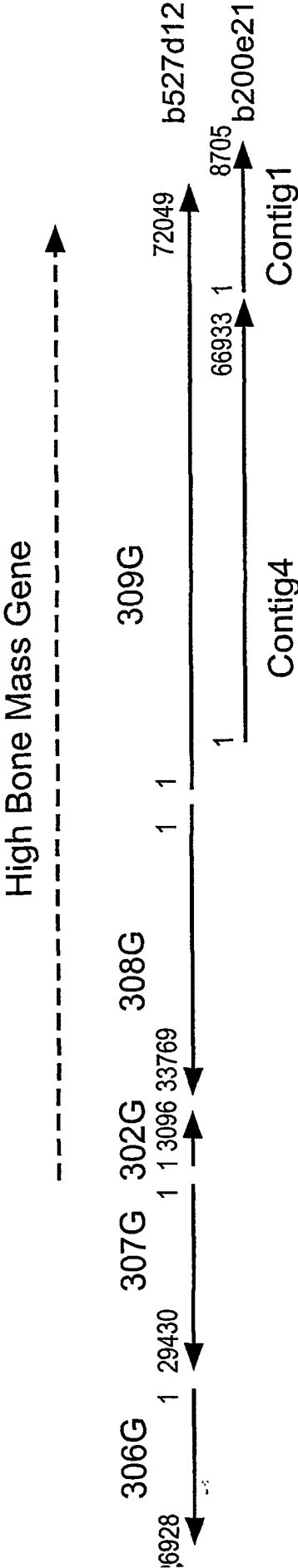


FIG. 5

13/89

FIG. 6A

1	ACTAAAGCGCCGCGCCATGGAGCCCGAGTGAGCGCGCGCGCGCGCGCCGTC	60
61	GCCGGACAACATGGAGGACGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
1	M E A A P P G P P W P L L L L L L L	17
121	GCTGCTGCTGGCGTGTGCGGTGCCCGGCCCGCGCGCGCGCGCGCGCGCGCT	180
18	L L L A L C G C P A P A A A S P L L L F	37
181	TGCCAACCGCGGACGTACGGCTGGTGACCGCGCGGAGTCAAGCTGGAGTCCACCAT	240
38	A N R R D V R L V D A G G V K L E S T I	57
241	CGTGGTCAGCGGCCCTGGAGGATGCGGCGCAGTGGACTTCCAGTTTCCAGGGAGCCGT	300
58	V V S G L E D A A A V D F Q F S K G A V	77
301	GTA CTGGACAGACGTGAGCGAGGAGGCCATCAAGCAGACCTACCTGAACCAAGCGGGGC	360
78	Y W T D V S E E A I K Q T Y L N Q T G A	97
361	CGCCGTGCAGAAACGTGGTCACTCCGGCCTGGTCTCTCCGACGGCCTCGCCTGCGACTG	420
98	A V Q N V V I S G L V S P D G L A C D W	117
421	GGTGGCAAGAAGCTGTACTGGACGGA CTGAGACCAACCGCATCGAGGTGGCCAACCT	480
118	V G K K L Y W T D S E T N R I E V A N L	137
481	CAATGGCACATCCCGGAAGGTGCTCTTCTGGCAGGACCTTGACCAGCCGAGGGCCATCGC	540
138	N G T S R K V L F W Q D L D Q P R A I A	157
541	CTTGACCCCGCTACGGGTACATGTACTGGACAGACTGGGGTGAGACGCCCGGATTGA	600
158	L D P A H G Y M Y W T D W G E T P R I E	177

14/89

FIG. 6B

601	GGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTTGTGACTCGGACATTTACTGGCC	660
178	R A G M D G S T R K I I V D S D I Y W P	197
661	CAATGGACTGACCATCGACCTGGAGGAGCAGAAAGCTCTACTGGGCTGACGCCAAGCTCAG	720
198	N G L T I D L E E Q K L Y W A D A K L S	217
721	CTTCATCCACCGTGCCAACCTGGACGGCTCGTTCGGCAGAAAGTGGTGGAGGGCAGCCT	780
218	F I H R A N L D G S F R Q K V V E G S L	237
781	GACGCACCCCTTCGCCCTGACGCTCTCCGGGGACACTCTGTACTGGACAGACTGGCAGAC	840
238	T H P F A L T L S G D T L Y W T D W Q T	257
841	CCGCTCCATCCATGCCTGCAACAAGCGCACTGGGGGAAGAGAGATCCTGAGTGC	900
258	R S I H A C N K R T G G K R K E I L S A	277
901	CCTCTACTCACCATGGACATCCAGGTGCTGAGCCAGGAGCGGAGCCTTCTTCCACAC	960
278	L Y S P M D I Q V L S Q E R Q P F F H T	297
961	TCGCTGTGAGGAGACAATGGCGGCTGCTCCACCTGTGCTGTCCCAAGCGAGCC	1020
298	R C E E D N G G C S H L C L L S P S E P	317
1021	TTTCTACATGCGCCTGCCCCACGGGTGTGCAGCTGCAGGACAACGGCAGGACGTGTAA	1080
318	F Y T C A C P T G V Q L Q D N G R T C K	337
1081	GGCAGGAGCCGAGGAGGTGCTGTGCTGGCCCCGGACGGACCTACGGAGGATCTCGCT	1140
338	A G A E E V L L L A R R T D L R R I S L	357

15/89

FIG. 6C

1141	GGACACGCCGGACTTCACCGACATCGTGTGACGATCCGGCAGCCATTGC	1200
358	D T P D F T D I V L Q V D D I R H A I A	377
1201	CATCGACTACGACCCGCTAGAGGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCAT	1260
378	I D Y D P L E G Y V Y W T D D E V R A I	397
1261	CCGACGGCGGTACCTGGACGGGTCTGGGGCGACGCTGGTCAACACCGAGATCAACGA	1320
398	R R A Y L D G S G A Q T L V N T E I N D	417
1321	CCCCGATGGCATCGCGGTCTGACTGGGTGGCCCGAAACCTCTACTGGACCGACACGGGCAC	1380
418	P D G I A V D W V A R N L Y W T D T G T	437
1381	GGACCGCATCGAGGTGACGCGCCTCAACGGCACCTCCCGCAAGATCCTGGTGTGCGAGGA	1440
438	D R I E V T R L N G T S R K I L V S E D	457
1441	CCTGGACGAGCCCCGAGCCATCGCACTGCACCCCGTGTGGGCCCTCATGTACTGGACAGA	1500
458	L D E P R A I A L H P V M G L M Y W T D	477
1501	CTGGGAGAGAACCTAAATCGAGTGTGCCAACTTGGATGGGCAGGAGCGCGGTGTGCT	1560
478	W G E N P K I E C A N L D G Q E R R V L	497
1561	GGTCAATGCCTCCCTCGGGTGGCCCAACGGCCCTGGCCCTTGACCTGCAGGAGGGGAAGCT	1620
498	V N A S L G W P N G L A L D L Q E G K L	517
1621	CTACTGGGAGACGCCAAGACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAG	1680
518	Y W G D A K T D K I E V I N V D G T K R	537

16/89

FIG. 6D

1681	GGGACCCCTCCTGGAGGACAAGCTCCCGCACATTTTCGGGTTTCACGCTGCTGGGGGACTT	1740
538	R T L L E D K L P H I F G F T L L G D F	557
1741	CATCTACTGGACTGACTGGCAGCGCCGCGAGCATCGAGCGGGTGCACAAGGTCAAGGCCAG	1800
558	I Y W T D W Q R R S I E R V H K V K A S	577
1801	CCGGGACGTCATCATTTGACCAAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGC	1860
578	R D V I I D Q L P D L M G L K A V N V A	597
1861	CAAGGTCGTCGGAACCAACCCGTGTGCGGACAGGAACGGGGTGCAGCCACCTGTGCTT	1920
598	K V V G T N P C A D R N G G C S H L C F	617
1921	CTTCACACCCACGCAACCCGGTGTGGCTGCCCCATCGGCCCTGGAGCTGCTGAGTGACAT	1980
618	F T P H A T R C G C P I G L E L L S D M	637
1981	GAAGACCTGCATCGTGCCCTGAGGCCCTTCTTGGTCTTCACCAAGCAGAGCCGCCATCCACAG	2040
638	K T C I V P E A F L V F T S R A A I H R	657
2041	GATCTCCCTCGAGACCAATAACAACGACGTGGCCATCCCGCTCACGGCGTCAAGGAGGC	2100
658	I S L E T N N N D V A I P L T G V K E A	677
2101	CTCAGCCCTGGACTTTGATGTGTCCAACAACCATCTACTGGACAGACGTCAAGCCTGAA	2160
678	S A L D F D V S N N H I Y W T D V S L K	697
2161	GACCATCAGCCGCGCCTTCATGAACGGGAGCTCGGTGGAGCACGTGGTGGAGTTTGGCCT	2220
698	T I S R A F M N G S S V E H V V E F G L	717

17/89

FIG. 6E

2221	TGACTACCCGAGGGCATGGCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCCGACAC	2280
718	D Y P E G M A V D W M G K N L Y W A D T	737
2281	TGGGACCAACAGAAATCGAAGTGGCGGCTGGACGGGCAGTTCCGGCAAGTCCTCGTGTG	2340
738	G T N R I E V A R L D G G Q F R Q V L V W	757
2341	GAGGACTTGGACAACCCGAGGTCGCTGGCCCTGGATCCCACCAAGGCTACATCTACTG	2400
758	R D L D N P R S L A L D P T K G Y I Y W	777
2401	GACCGAGTGGGGCGGCAAGCCGAGGATCGTGCGGGCCTTCATGACGGGACCAACTGCAT	2460
778	T E W G G K P R I V R A F M D G T N C M	797
2461	GACGCTGGTGGACAAGGTGGGCCGGGCCAACGACCTCACCATTTGACTACGCTGACCAGCG	2520
798	T L V D K V G R A N D L T I D Y A D Q R	817
2521	CCTCTACTGGACCGACCTGGACACCAACATGATCGAGTCGTCCAACATGCTGGGTCAGGA	2580
818	L Y W T D L D T N M I E S S N M L G Q E	837
2581	GCGGGTCGTGATTGCCGACGATCTCCCGCACCCCGTTCCGGTCTGACGCAGTACAGCGATTA	2640
838	R V V I A D D L P H P F G L T Q Y S D Y	857
2641	TATCTACTGGACAGACTGGAATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCGGCCG	2700
858	I Y W T D W N L H S I E R A D K T S G R	877
2701	GAACCGCACCCCTCATCCAGGGCCACCTGGACTTCGTGATGGACATCCTGGTGTTCCTCACTC	2760
878	N R T L I Q G H L D F V M D I L V F H S	897

18/89

FIG. 6F

2761	CTCCGCCAGGATGGCCTCAATGACTGTATGCACAACAACGGGAGTGTGGGCAGCTGTG	2820
898	S R Q D G L N D C M H N N G Q C G Q L C	917
2821	CCTTGCCATCCCCGGCGGCCACCGCTGGCGGCTGGCCTCACACTACACCCCTGGACCCAG	2880
918	L A I P G G H R C G C A S H Y T L D P S	937
2881	CAGCCGCAACTGCAGCCCGCCACACCTTCTTGCTGTTCAGCCAGAAATCTGCCATCAG	2940
938	S R N C S P P T T F L L F S Q K S A I S	957
2941	TCGGATGATCCCCGGACGACGACAGCCCCGGATCTCATCTGCCCTGCTGGACTGAG	3000
958	R M I P D D Q H S P D L I L P L H G L R	977
3001	GAACGTCAAAGCCATCGACTATGACCCACTGGACAAGTTCACTACTGGTGGATGGCGG	3060
978	N V K A I D Y D P L D K F I Y W V D G R	997
3061	CCAGAACATCAAGCGAGCCAAGGACGACGGACCCAGCCCTTTGTGTTTGACCTCTCTGAG	3120
998	Q N I K R A K D D G T Q P F V L T S L S	1017
3121	CCAAGGCCAAACCCAGACAGGCAGCCCCACGACCTCAGCATCGACATCTACAGCCGAC	3180
1018	Q G Q N P D R Q P H D L S I D I Y S R T	1037
3181	ACTGTTCTGGACGTGCGAGGCCACCAATACCATCAACGTCCACAGGCTGAGCGGGAAGC	3240
1038	L F W T C E A T N T I N V H R L S G E A	1057
3241	CATGGGGGTGCTGCGTGGGACCGCGACAAGCCAGGCCATCGTCGTCAACGCGGA	3300
1058	M G V V L R G D R D K P R A I V V N A E	1077

19/89

FIG. 6G

3301	GCGAGGGTACCTGTACTTCACCAACATGCAGGACCGGGCAGCCAAGATCGAAACGCGCAGC	3360
1078	R G Y L Y F T N M Q D R A A K I E R A A	1097
3361	CCTGGACGGCACCGAGCGAGGTCCTCTTCACCACCGGCCCTCATCCGCCCTGTGGCCCT	3420
1098	L D G T E R E V L F T T G L I R P V A L	1117
3421	GGTGGTGACAACACACTGGGCAAGCTGTTCTGGGTGACGCGGACCTGAAGCGCATTTGA	3480
1118	V V D N T L G K L F W V D A D L K R I E	1137
3481	GAGCTGTGACCTGTACGGGCCAACCGCCTGACCCCTGGAGGACGCCAACATCGTGCAGCC	3540
1138	S C D L S G A N R L T L E D A N I V Q P	1157
3541	TCTGGGCCCTGACCATCCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCAGCATGAT	3600
1158	L G L T I L G K H L Y W I D R Q Q M I	1177
3601	CGAGCGTGTGGAGAAGACCCGGGGACAAGCGGACTCGCATCCAGGGCCGTGTCGCCCA	3660
1178	E R V E K T T G D K R T R I Q G R V A H	1197
3661	CCTCACTGGCATCCATGCAGTGGAGGAAGTCAGCCTGGAGGAGTTCTCAGCCCACCCATG	3720
1198	L T G I H A V E E V S L E E F S A H P C	1217
3721	TGCCCCGTGACAAATGGTGGCTGCTCCACATCTGTATTGCCAAGGGTGATGGGACACCCAG	3780
1218	A R D N G G C S H I C I A K G D G T P R	1237
3781	GTGCTCATGCCCAGTCCACCTCGTGTCTCTGCAGAACCTGTGACCTGTGGAGAGCCGCC	3840
1238	C S C P V H L V L L Q N L L T C G E P P	1257

20/89

FIG. 6H

3841	CACCTGCTCCCGGACCAAGTTTGCAATGTCACAGGGAGATCGACTGTATCCCGGGGC	3900
1258	T C S P D Q F A C A T G E I D C I P G A	1277
3901	CTGGCGCTGTGACGGCTTTCCCGAGTGCATGACCAAGACGAGAGGGCTGCCCGGT	3960
1278	W R C D G F P E C D D Q S D E E G C P V	1297
3961	GTGCTCCGCGCCAGTTCCCTCGCGGGGTCAAGTGTGTGGACCTGCGCCTGCGCTG	4020
1298	C S A A Q F P C A R G Q C V D L R L R C	1317
4021	CGACGGCGAGGCACTGTCAAGACCGCTCAGACGAGGTGACTGTGACGCCATCTGCCT	4080
1318	D G E A D C Q D R S D E V D C D A I C L	1337
4081	GCCCAACCAAGTTCCGGTGTGCGAGCGGCCAGTGTGTCTCATCAACAGCAGTGCGACTC	4140
1338	P N Q F R C A S G Q C V L I K Q Q C D S	1357
4141	CTTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGTGAATCACCAAGCCGCCCTC	4200
1358	F P D C I D G S D E L M C E I T K P P S	1377
4201	AGACGACAGCCCGCCACAGCAGTGCCATCGGGCCCGTCAATGGCATCATCCTCTCTCT	4260
1378	D D S P A H S S A I G P V I G I I L S L	1397
4261	CTTCGTCAATGGTGTCTATTTTGTGTGCCAGCGCGTGTGTGCCAGCGCTATGCGGG	4320
1398	F V M G G V Y F V C Q R V V C Q R Y A G	1417
4321	GGCCAACGGGCCCTTCCCGCACGAGTATGTACGGGGACCCCGCACGTGCCCTCAATT	4380
1418	A N G P F P H E Y V S G T P H V P L N F	1437

21/89

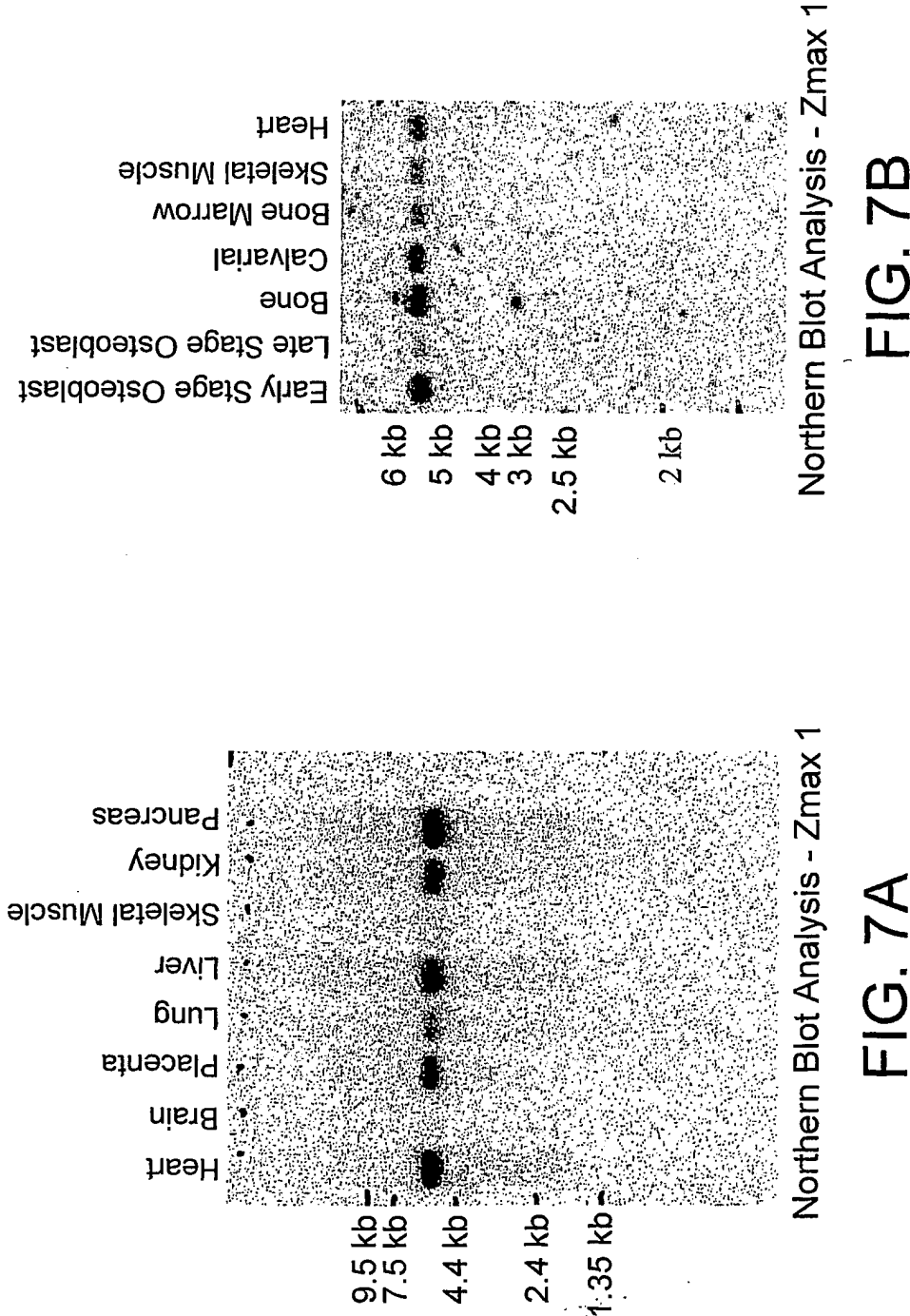
FIG. 6I

4381	CATAGCCCCGGGGTTCCAGCATGGCCCCCTTCACAGGCATCGCGAAAGTCCAT	4440
1438	I A P G G S Q H G P F T G I A C G K S M	1457
4441	GATGAGCTCCGTGAGCCCTGATGGGGGGCGGGGGTGTGCCCCCTCTACGACCGGAACCA	4500
1458	M S S V S L M G G R G G V P L Y D R N H	1477
4501	CGTCACAGGGCCTCGTCCAGCAGCTCGTCCAGCACGAAGGCCACGCTGTACCCGCCGAT	4560
1478	V T G A S S S S S S T K A T L Y P P I	1497
4561	CCTGAACCCGGCCCTCCCGGCCACGGACCCCTCCCTGTACAACATGGACATGTTCTA	4620
1498	L N P P P S P A T D P S L Y N M D M F Y	1517
4621	CTCTTCAAACATTCCGGCCACTGCGAGACCGTACAGGCCCTACATCATTCGAGGAATGGC	4680
1518	S S N I P A T A R P Y R P Y I I R G M A	1537
4681	GCCCCGACGACGCCCTGCAGCACCGACGTGTGTGACAGCGACTACAGCGCCAGCCGCTG	4740
1538	P P T T P C S T D V C D S D Y S A S R W	1557
4741	GAAGGCCAGCAAGTACTACCTGGATTGAACTCGGACTCAGACCCCTATCCACCCACC	4800
1558	K A S K Y Y L D L N S D S D P Y P P P	1577
4801	CACGCCCCACAGCCAGTACCTGTGCGGGGAGGACAGCTGCCCGCCCTCGCCCGCCACCGA	4860
1578	T P H S Q Y L S A E D S C P P S P A T E	1597
4861	GAGGAGCTACTTCCATCTCTTCCCGCCCCCTCCGTCCCCCTGCACGGACTCATCTGACC	4920
1598	R S Y F H L F P P P P S P C T D S S	1615

22/89

FIG. 6J

4921	TCGGCCGGGCCACTCTGGCTTCTCTGTGCCCCCTGTAAATAGTTTAAATATGAACAAAGA	4980
4981	AAAAAATATATTTTATGATTTTAAATAATAATAATAATGCGATTTTAAATAACATGAGAAA	5040
5041	TGTGAACCTGTGATGGGTGGCAGGGCTGGGAGAACTTTGTACAGTGGAGAAATATTAT	5100
5101	AAACTTAATTTTGTAAACA	5120



Zmax 1 random samples

b527d12-h_Contig087C_1.nt

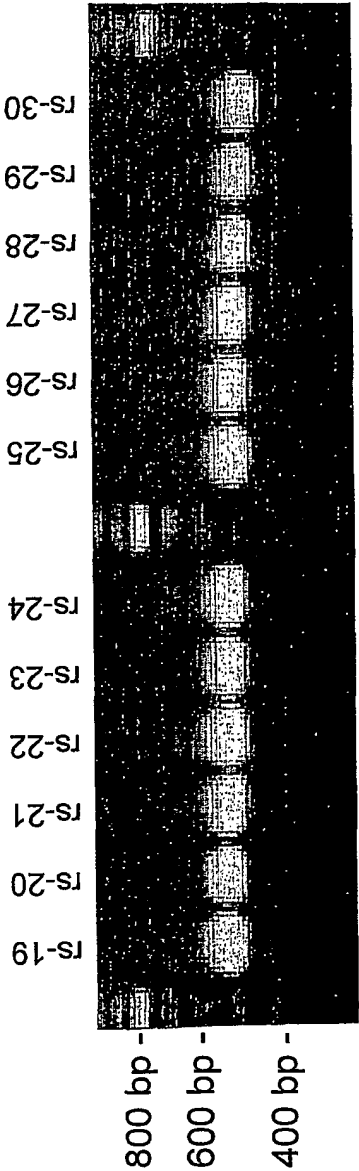


FIG. 8

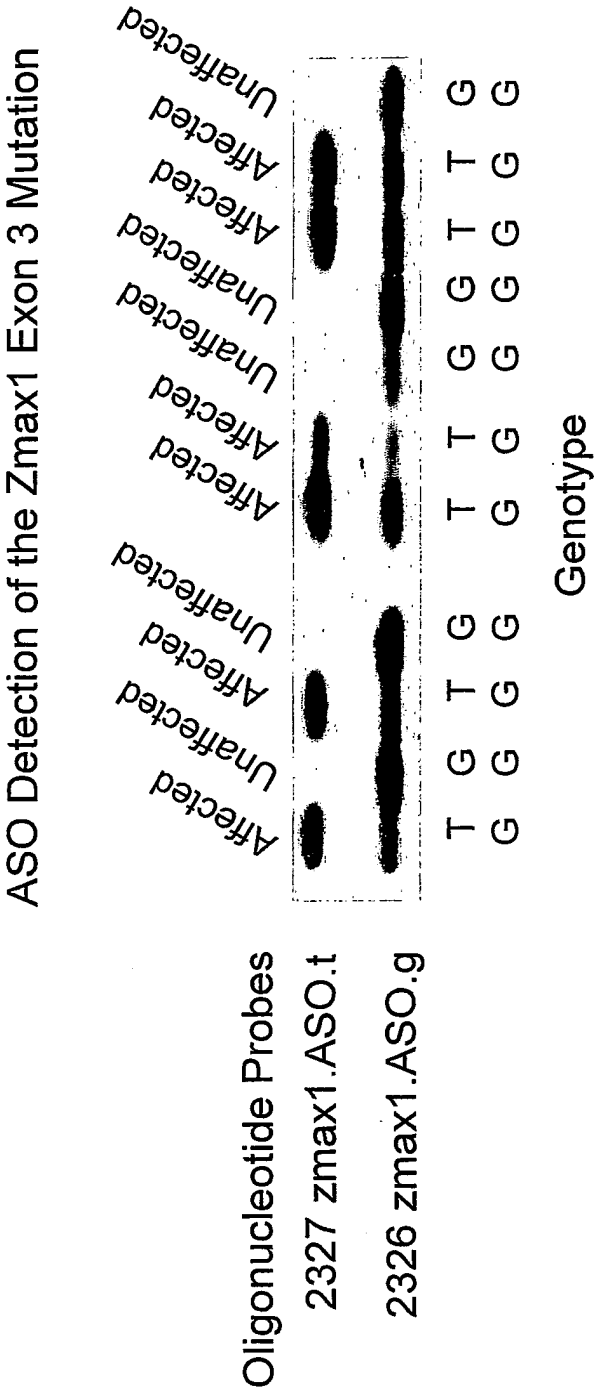


FIG. 9

26/89

Mouse Zmax1 In situ hybridization
100X Magnification

Antisense probe



FIG. 10A

Mouse Zmax1 In situ hybridization
100X Magnification

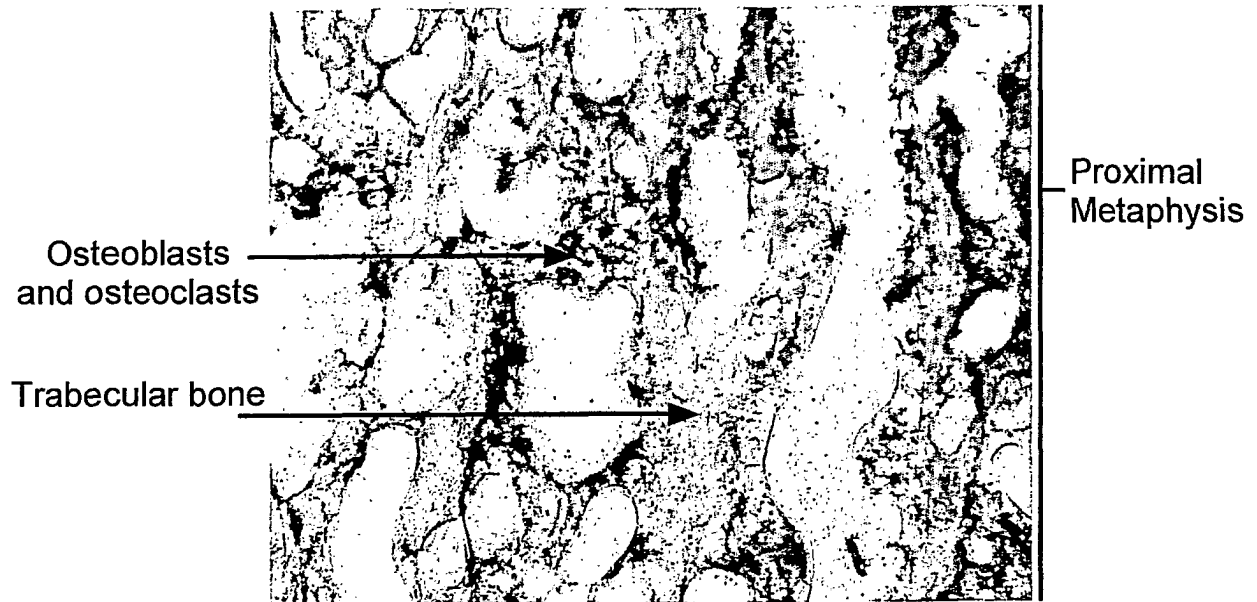
Sense probe



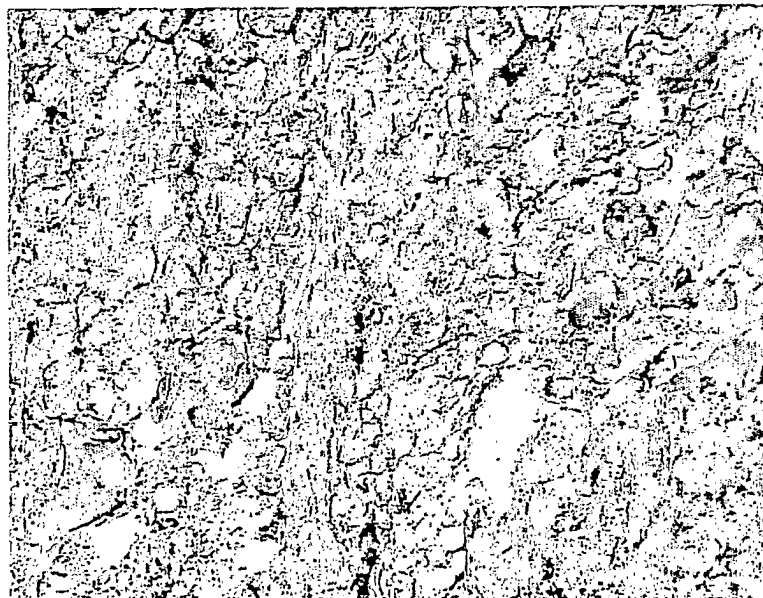
FIG. 10B

27/89

Mouse Zmax1 In situ hybridization
400X Magnification
Antisense probe

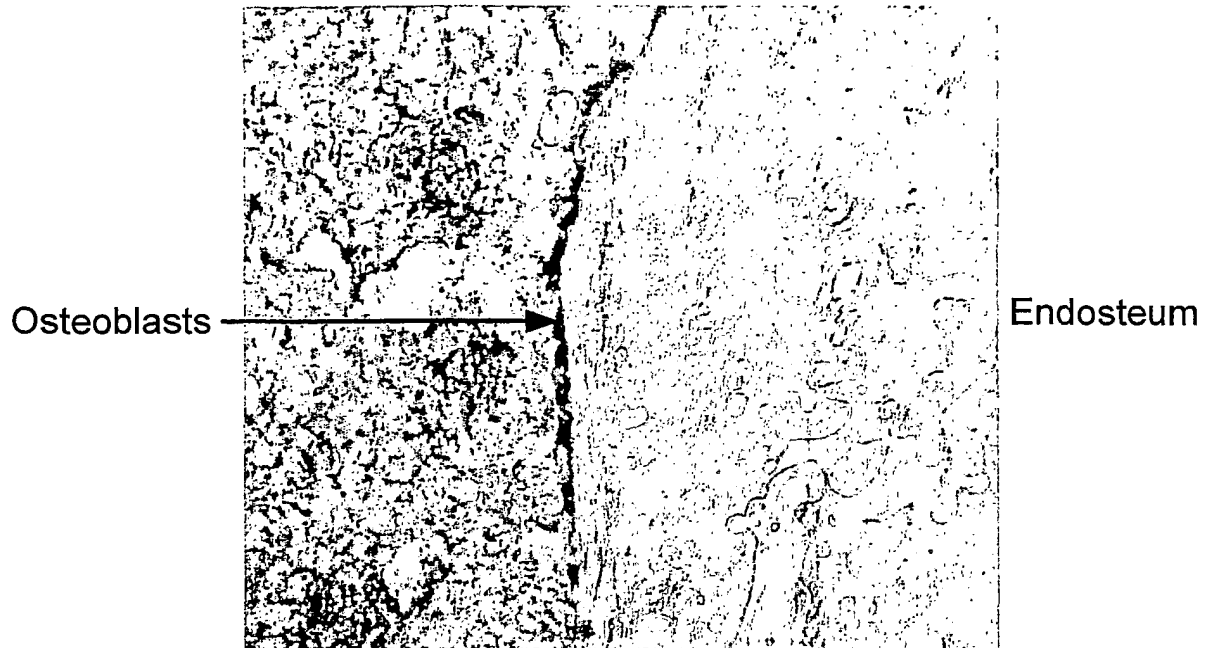
**FIG. 11A**

Mouse Zmax1 In situ hybridization
400X Magnification
Sense probe

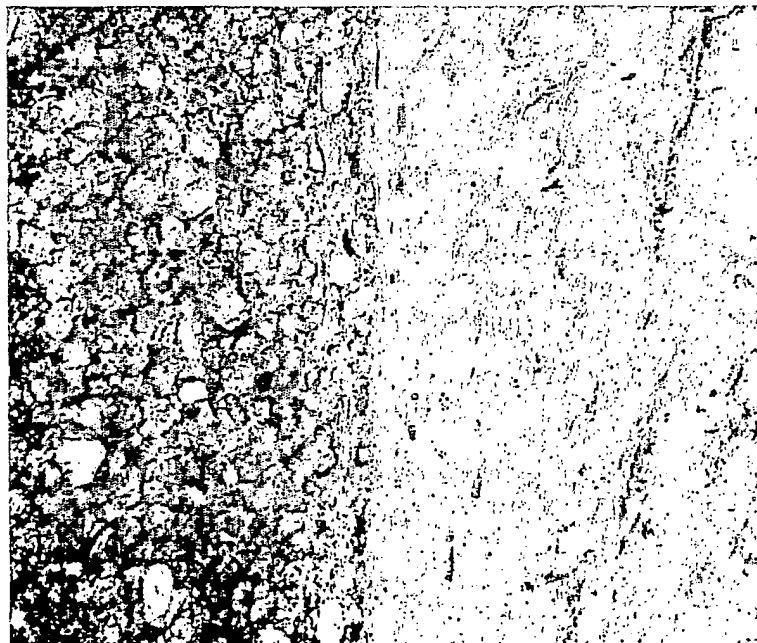
**FIG. 11B**

28/89

Mouse Zmax1 In situ hybridization
400X Magnification
Antisense probe

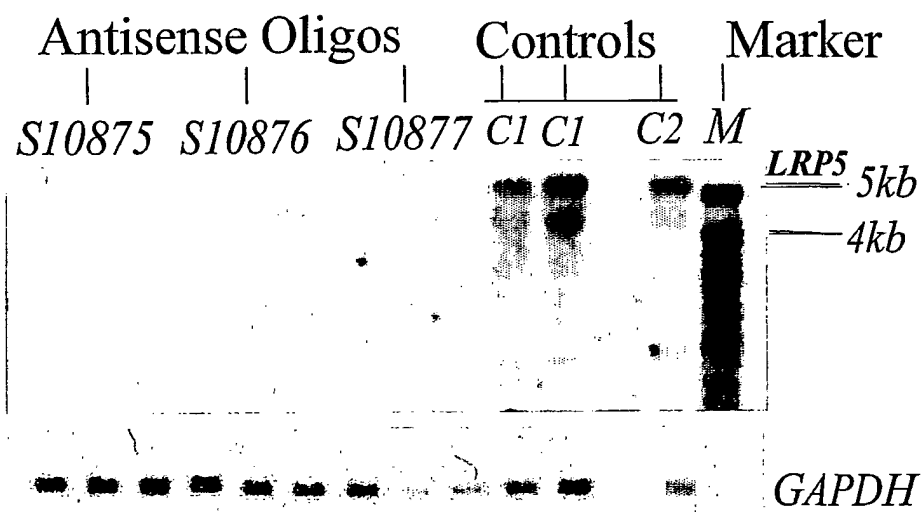
**FIG. 12A**

Mouse Zmax1 In situ hybridization
400X Magnification
Sense probe

**FIG. 12B**

29/89

Antisense Inhibition of Zmax1 Expression



MC-3T3 cells

FIG. 13

Zmax1 Exon3 ASO Assay

T-specific Oligo
58 °C Wash

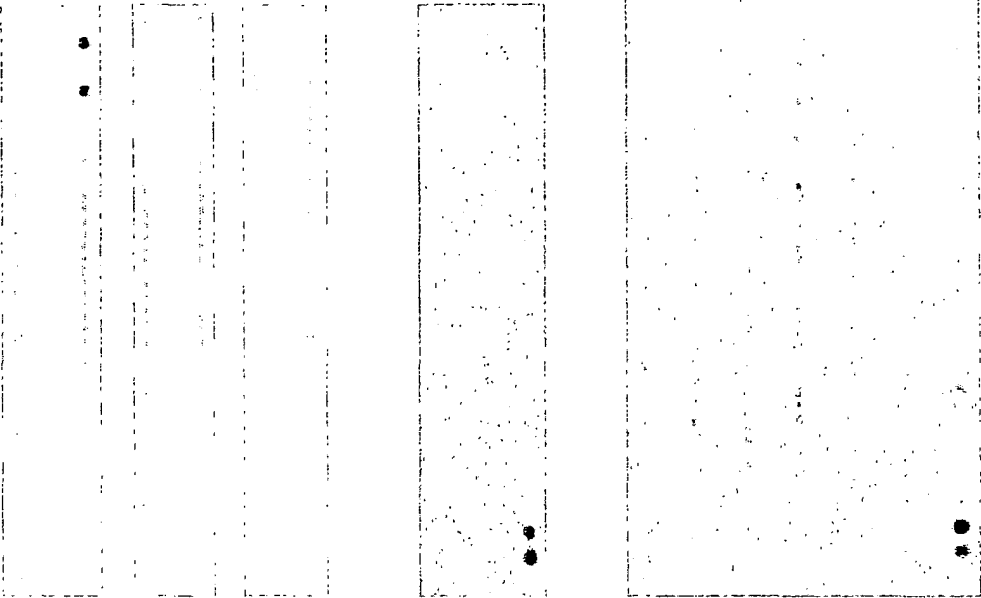


FIG. 14B

Zmax1 Exon3 ASO Assay

G-specific Oligo
55 °C Wash

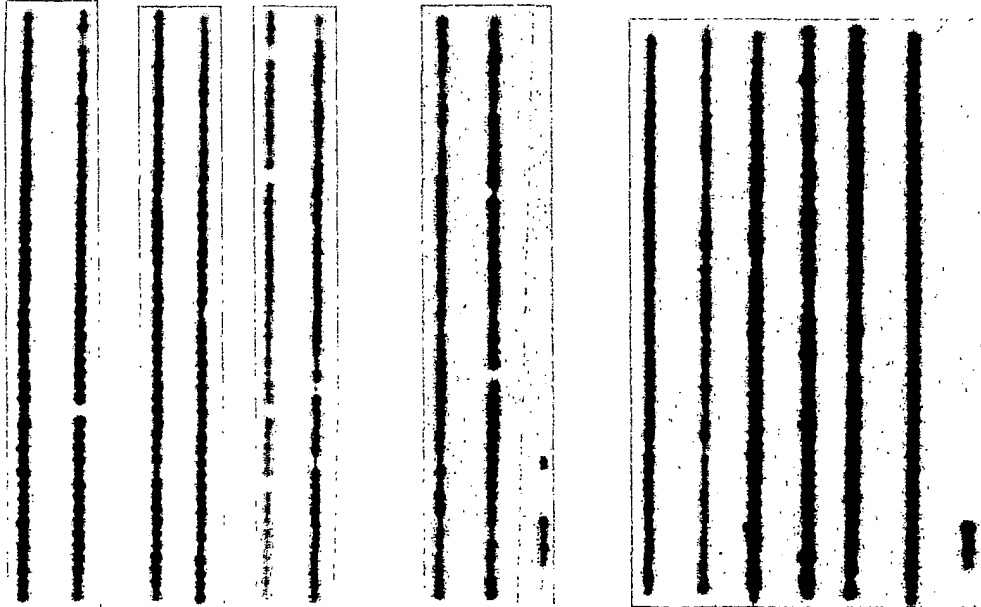


FIG. 14A

Possible Involvement of Zmax1 in Focal Adhesion Signaling

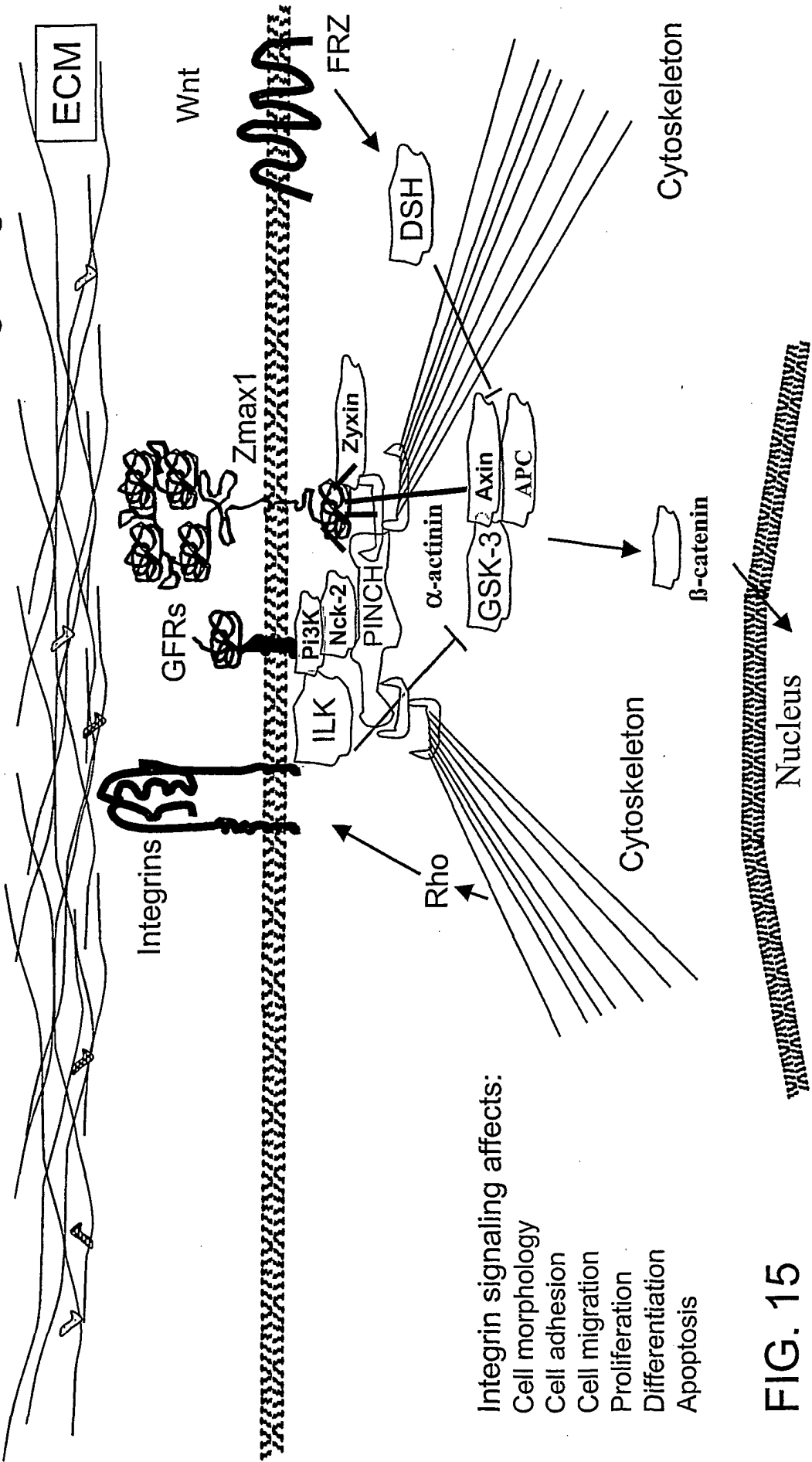


FIG. 15

ZMAX Gene Targeting Knock-in / Knock-out vectors

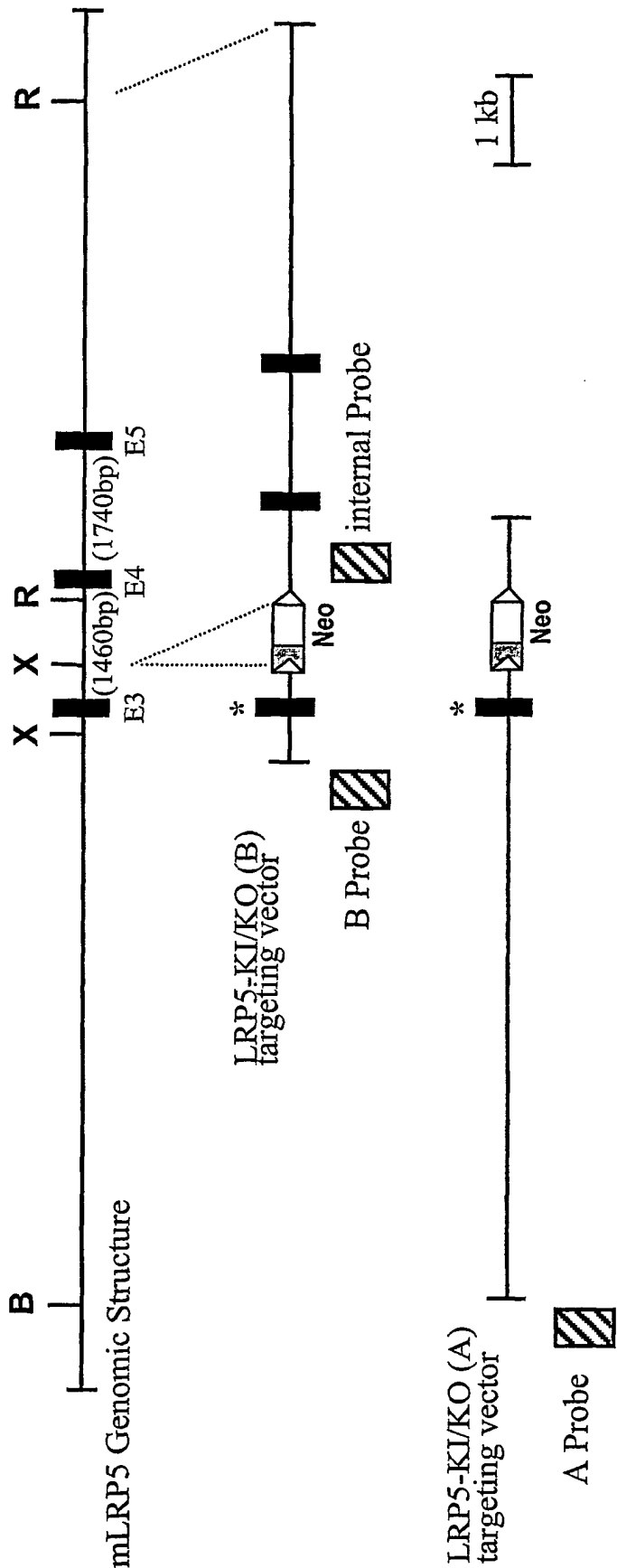


FIG. 16

33/89

HBM: Constructs for transgenic mice

*Confirmation of expression**Transient transfection into HOB-02-02 cells**

CMV β Actin		Type I collagen	
HBM	Zmax1	HBM	Zmax1
X 1,000	X 1,000	X 10	X 10

**Fold increase compared to Zmax1
in cells transfected with empty vector*

FIG. 17

34/89

Comparison of Human & Mouse TaqMan
Primer/Probe Sets for Zmax1/HBM

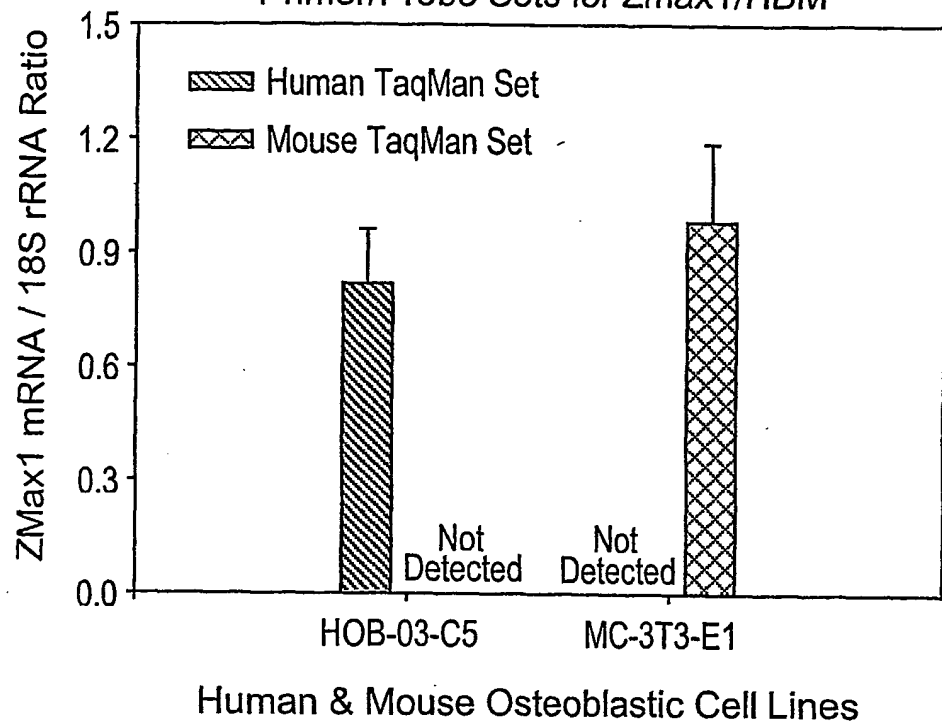


FIG. 18

Quantification of Human Zmax1 mRNA in a
Mixed Human & Mouse RNA Background

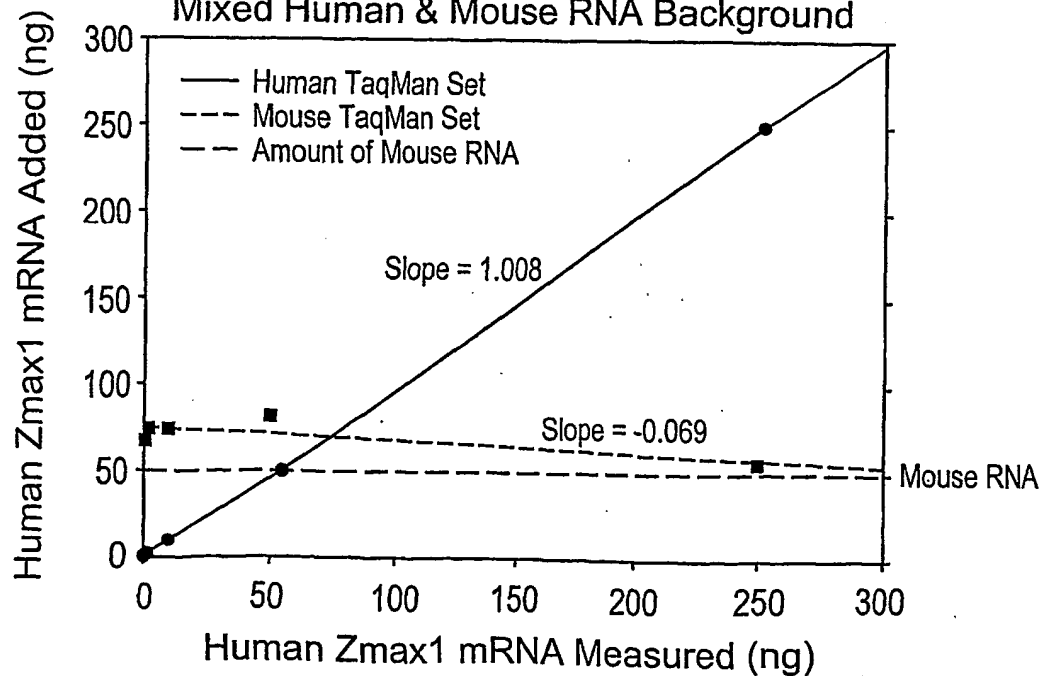


FIG. 19

35/89

HBM: Transgenic mice
*mRNA expression by Taqman analysis**

Line	Tissue						
	Tibia	Femur	Heart	Gonad	Brain	Kidney	Liver
HBMMCBA							
2	7-10		20-90	2-30	6-11	5-9	<1
13	1-2		6-7	3-4	5-6	<1	<1
18	10-11						
HBMMTIC							
19	7-8	19-20	1				1
35	1	1	0				0

* relative to Zmax1 in HOB-03-C5

FIG. 20

HBM: In vivo pDXA - transgenic mice
HBMMCBA construct

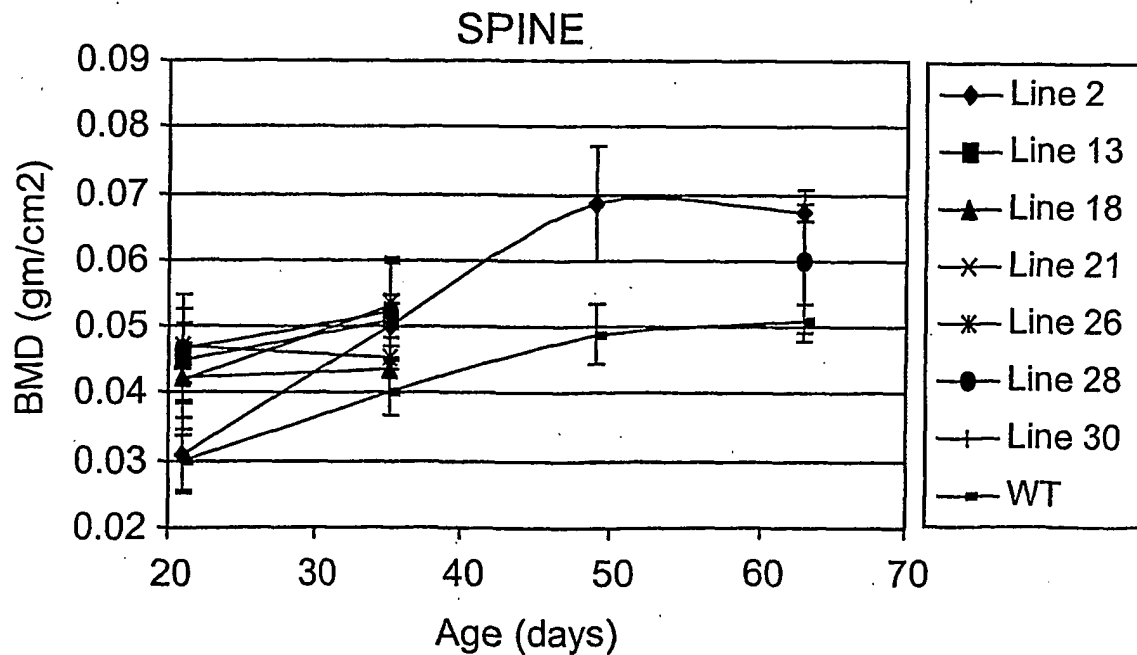


FIG. 21A

36/89

HBM: In vivo pDXA - transgenic mice

HBMMCBA construct

FEMUR

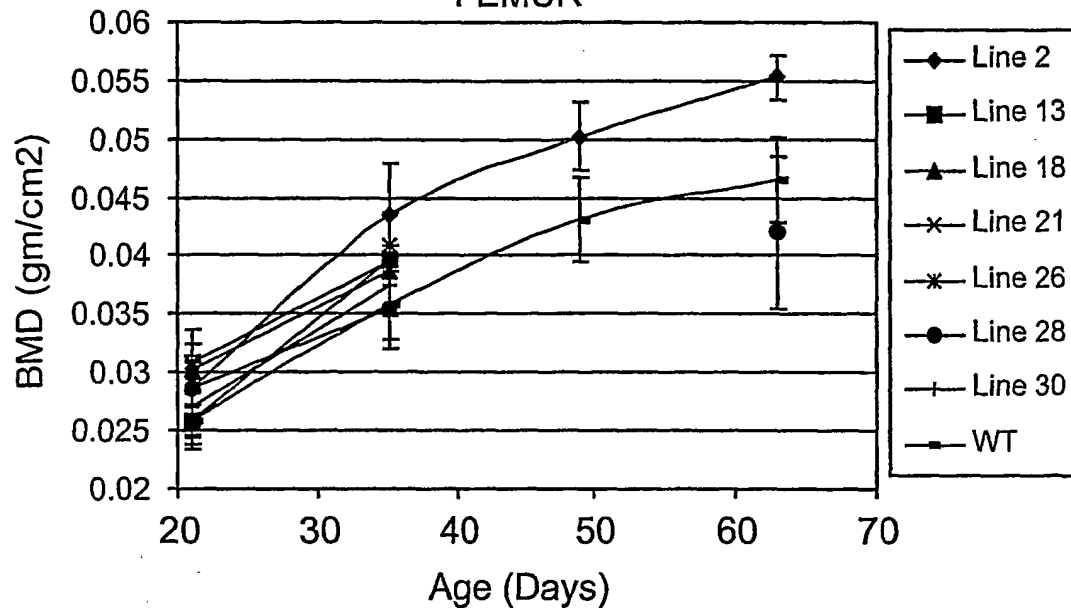


FIG. 21B

HBM: In vivo pDXA - transgenic mice

HBMMCBA construct

T-TOTAL

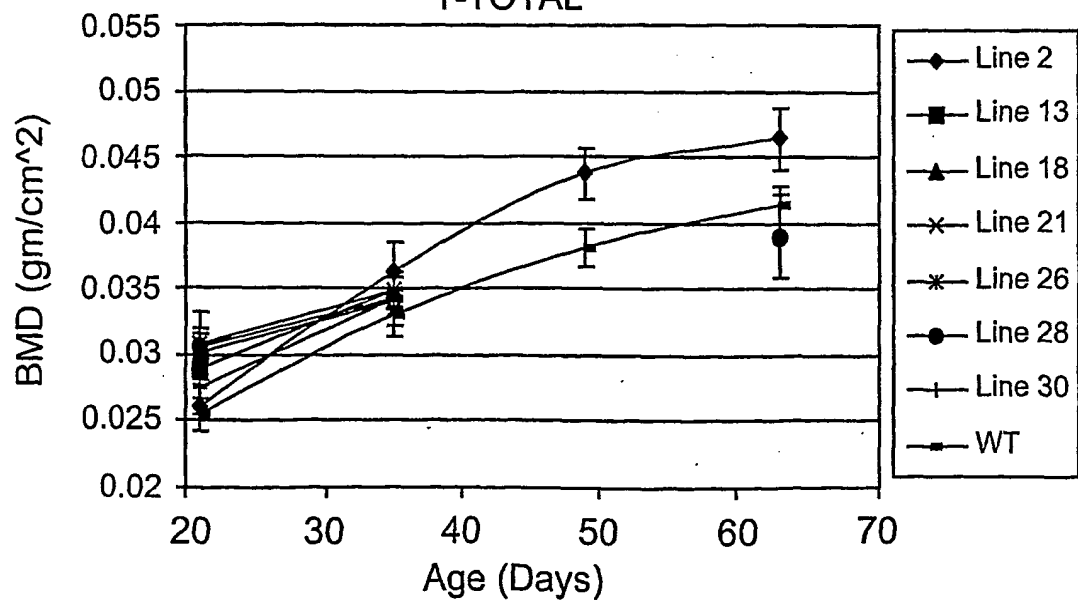


FIG. 21C

37/89

HBM: In vivo pDXA - transgenic mice

HBMMTIC construct

SPINE

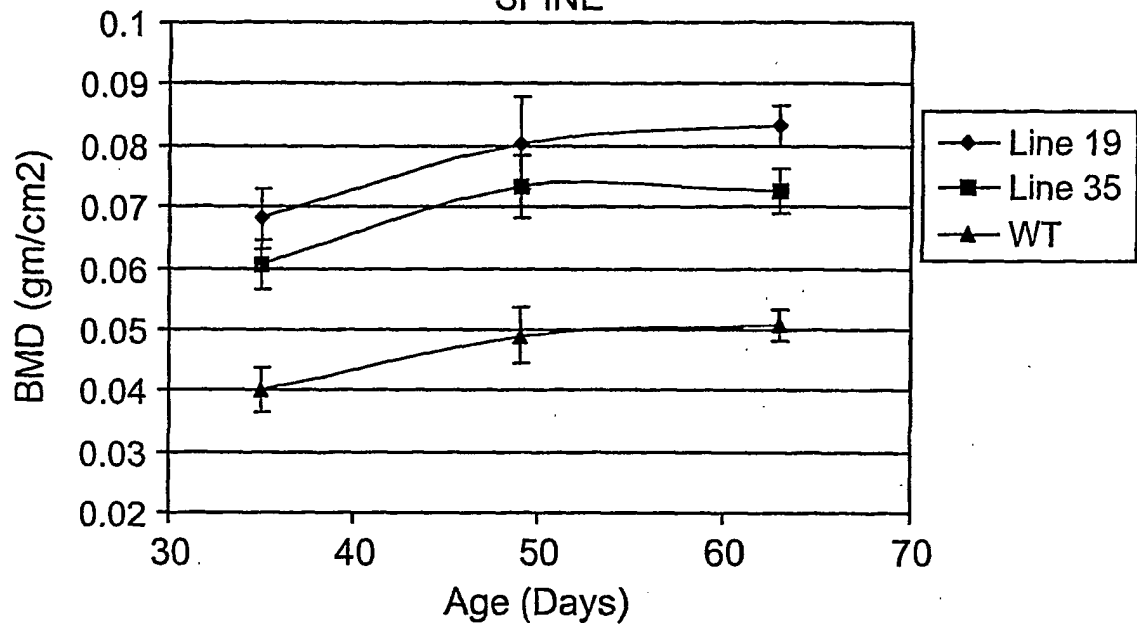


FIG. 21D

HBM: In vivo pDXA - transgenic mice

HBMMTIC construct

FEMUR

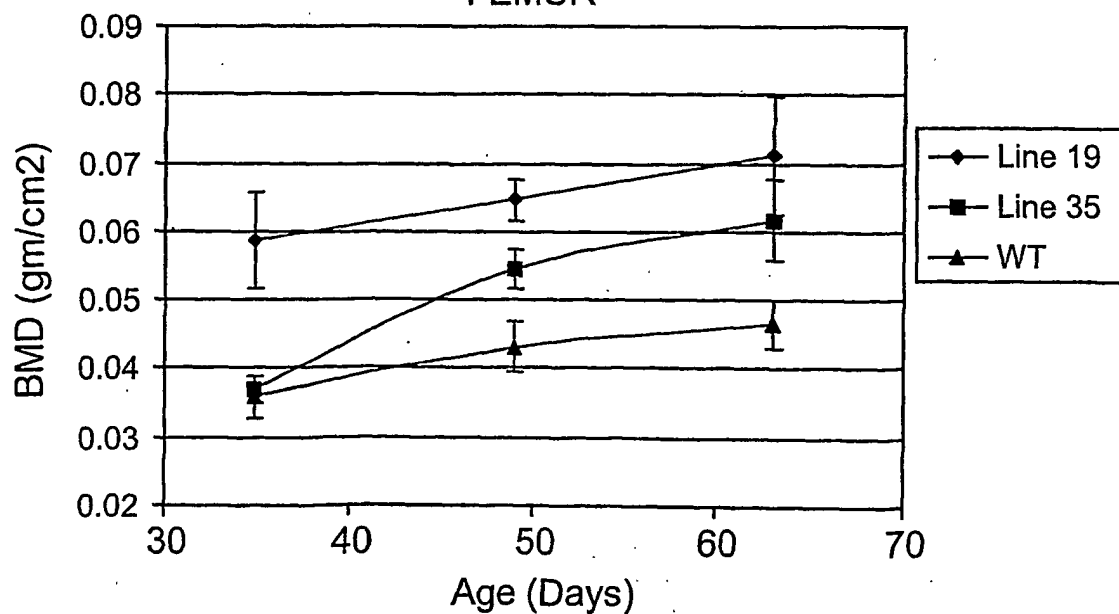


FIG. 21E

38/89

HBM: In vivo pDXA - transgenic mice

HBMMTIC construct

T-TOTAL

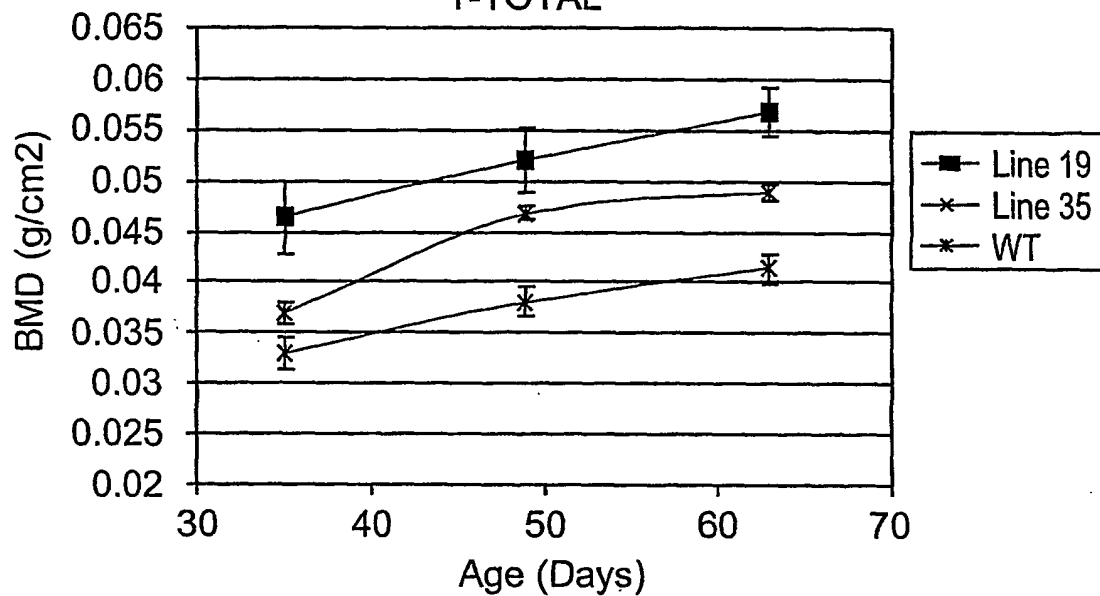


FIG. 21F

39/89

HBM transgenic mice: in vivo pDXA*
BMD % changes vs WT in 5 week old animals

Line	n	Femur	Spine	Total	Tibia mRNA
HBMMCBA					
2	11	21	24	10	x7-10
13	4	11	27	5	x1-2
18	1	8	9	3	x10-11
21	1	10	12	6	0
28	15	0	30	4	x1
30	7	4	32	4	0
HBMMTIC					
19	5	63	70	41	x7-8
35	1	4	47	6	x1

* CV: Femur (2.7%); Spine (6.4%); Total (1.7%)

FIG. 22

HBM transgenic mice: in vivo pDXA*
BMD % changes vs WT in 9 week old animals

Line	n	Femur	Spine	Total
HBMMCBA				
2	3	19	32	12
HBMMTIC				
19	2	52	64	37
35	3	32	43	18

* CV: Femur (2.7%); Spine (6.4%); Total (1.7%)

FIG. 23

40/89

FIG. 24A

HBMGI_2AS

TCTAGACTCG AGGCGGCCGC CCATTGTGCA CTAAAGCGCC GCCGCCGCGC CATGGAGCCC
GAGTGAGCGC GGC GCGGGCC CGTCCGGCCG CCGGACAACA TGGAGGCAGC GCCGCCCGGG
CCGCCGTGGC CGCTGCTGCT GCTGCTGCTG CTGCTGCTGG CGCTGTGCGG CTGCCCCGGC
CCCCCGCGG CCTCGCCGCT CCTGCTATTT GCCAACC GCCGACGTACG GCTGGTGGAC
GCCGGCGGAG TCAAGCTGGA GTCCACCATC GTGGTCAGCG GCCTGGAGGA TGC GGCCGCA
GTGGACTTCC AGTTTTCCAA GGGAGCCGTG TACTGGACAG ACGTGAGCGA GGAGGCCATC
AAGCAGACCT ACCTGAACCA GACGGGGGCC GCCGTGCAGA ACGTGGTCAT CTCCGGCCTG
GTCTCTCCCG ACGGCCTCGC CTGCGACTGG GTGGGCAAGA AGCTGTACTG GACGGACTCA
GAGACCAACC GCATCGAGGT GGCCAACCTC AATGGCACAT CCCGGAAGGT GCTCTTCTGG
CAGGACCTTG ACCAGCCGAG GGCCATCGCC TTGGACCCCG CTCACGGGTA CATGTACTGG
ACAGACTGGG TTGAGACGCC CCGGATTGAG CGGGCAGGGA TGGATGGCAG CACCCGGAAG
ATCATTGTGG ACTCGGACAT TTA CTGGCCC AATGGACTGA CCATCGACCT GGAGGAGCAG
AAGCTCTACT GGGCTGACGC CAAGCTCAGC TTCATCCACC GTGCCAACCT GGACGGCTCG
TTCCGGCAGA AGGTGGTGGA GGGCAGCCTG ACGCACCCCT TCGCCCTGAC GCTCTCCGGG
GACACTCTGT ACTGGACAGA CTGGCAGACC CGCTCCATCC ATGCCTGCAA CAAGCGCACT
GGGGGGAAGA GGAAGGAGAT CCTGAGTGCC CTCTACTCAC CCATGGACAT CCAGGTGCTG
AGCCAGGAGC GGCAGCCTTT CTTCCACACT CGCTGTGAGG AGGACAATGG CGGCTGCTCC
CACCTGTGCC TGCTGTCCCC AAGCGAGCCT TTCTACACAT GCGCCTGCCC CACGGGTGTG
CAGCTGCAGG ACAACGGCAG GACGTGTAAG GCAGGAGCCG AGGAGGTGCT GCTGCTGGCC
CGGCGGACGG ACCTACGGAG GATCTCGCTG GACACGCCGG ACTTCACCGA CATCGTGCTG
CAGGTGGACG ACATCCGGCA CGCCATTGCC ATCGACTACG ACCCGCTAGA GGGCTATGTC
TACTGGACAG ATGACGAGGT GCGGGCCATC CGCAGGGCGT ACCTGGACGG GTCTGGGGCG
CAGACGCTGG TCAACACCGA GATCAACGAC CCCGATGGCA TCGCGGTCGA CTGGGTGGCC
CGAAACCTCT ACTGGACCGA CACGGGCACG GACCGCATCG AGGTGACGCG CCTCAACGGC

41/89

FIG. 24B

HBMGI_2AS

ACCTCCCGCA AGATCCTGGT GTCGGAGGAC CTGGACGAGC CCCGAGCCAT CGCACTGCAC
CCCGTGATGG GCCTCATGTA CTGGACAGAC TGGGGAGAGA ACCCTAAAAT CGAGTGTGCC
AACTTGGATG GGCAGGAGCG GCGTGTGCTG GTCAATGCCT CCCTCGGGTG GCCCAACGGC
CTGGCCCTGG ACCTGCAGGA GGGGAAGCTC TACTGGGGAG ACGCCAAGAC AGACAAGATC
GAGGTGATCA ATGTTGATGG GACGAAGAGG CGGACCCTCC TGGAGGACAA GCTCCCGCAC
ATTTTCGGGT TCACGCTGCT GGGGGACTTC ATCTACTGGA CTGACTGGCA GCGCCGCAGC
ATCGAGCGGG TGCACAAGGT CAAGGCCAGC CGGGACGTCA TCATTGACCA GCTGCCCCGAC
CTGATGGGGC TCAAAGCTGT GAATGTGGCC AAGGTCGTCTG GAACCAACCC GTGTGCGGAC
AGGAACGGGG GGTGCAGCCA CCTGTGCTTC TTCACACCCC ACGCAACCCG GTGTGGCTGC
CCCATCGGCC TGGAGCTGCT GAGTGACATG AAGACCTGCA TCGTGCCTGA GGCCTTCTTG
GTCTTCACCA GCAGAGCCGC CATCCACAGG ATCTCCCTCG AGACCAATAA CAACGACGTG
GCCATCCCGC TCACGGGCGT CAAGGAGGCC TCAGCCCTGG ACTTTGATGT GTCCAACAAC
CACATCTACT GGACAGACGT CAGCCTGAAG ACCATCAGCC GCGCCTTCAT GAACGGGAGC
TCGGTGGAGC ACGTGGTGGG GTTTGGCCTT GACTACCCCG AGGGCATGGC CGTTGACTGG
ATGGGCAAGA ACCTCTACTG GGCCGACACT GGGACCAACA GAATCGAAGT GGCGCGGCTG
GACGGGCAGT TCCGGCAAGT CCTCGTGTGG AGGGACTTGG ACAACCCGAG GTCGCTGGCC
CTGGATCCCA CCAAGGGCTA CATCTACTGG ACCGAGTGGG GCGGCAAGCC GAGGATCGTG
CGGGCCTTCA TGGACGGGAC CAACTGCATG ACGCTGGTGG ACAAGGTGGG CCGGGCCAAC
GACCTCACCA TTGACTACGC TGACCAGCGC CTCTACTGGA CCGACCTGGA CACCAACATG
ATCGAGTCGT CCAACATGCT GGGTCAGGAG CGGGTCGTGA TTGCCGACGA TCTCCCGCAC
CCGTTTCGGT TGACGCAGTA CAGCGATTAT ATCTACTGGA CAGACTGGAA TCTGCACAGC
ATTGAGCGGG CCGACAAGAC TAGCGGCCGG AACC GCACCC TCATCCAGGG CCACCTGGAC
TTCGTGATGG ACATCCTGGT GTTCCACTCC TCCCGCCAGG ATGGCCTCAA TGACTGTATG

42/89

FIG. 24C

HBMGI_2AS

CACAACAACG GGCAGTGTGG GCAGCTGTGC CTTGCCATCC CCGGCGGCCA CCGCTGCGGC
TGCGCCTCAC ACTACACCCT GGACCCCAGC AGCCGCAACT GCAGCCCGCC CACCACCTTC
TTGCTGTTCA GCCAGAAATC TGCCATCAGT CGGATGATCC CGGACGACCA GCACAGCCCG
GATCTCATCC TGCCCCTGCA TGGACTGAGG AACGTCAAAG CCATCGACTA TGACCCACTG
GACAAGTTCA TCTACTGGGT GGATGGGCGC CAGAACATCA AGCGAGCCAA GGACGACGGG
ACCCAGCCCT TTGTTTTGAC CTCTCTGAGC CAAGGCCAAA ACCCAGACAG GCAGCCCCAC
GACCTCAGCA TCGACATCTA CAGCCGGACA CTGTTCTGGA CGTGCGAGGC CACCAATACC
ATCAACGTCC ACAGGCTGAG CGGGGAAGCC ATGGGGGTGG TGCTGCGTGG GGACCGCGAC
AAGCCCAGGG CCATCGTCGT CAACGCGGAG CGAGGGTACC TGTACTTCAC CAACATGCAG
GACCGGGCAG CCAAGATCGA ACGCGCAGCC CTGGACGGCA CCGAGCGCGA GGTCTCTTC
ACCACCGGCC TCATCCGCCC TGTGGCCCTG GTGGTAGACA ACACACTGGG CAAGCTGTTC
TGGGTGGACG CGGACCTGAA GCGCATTGAG AGCTGTGACC TGTCAGGGGC CAACCGCCTG
ACCCTGGAGG ACGCCAACAT CGTGCAGCCT CTGGGCCTGA CCATCCTTGG CAAGCATCTC
TACTGGATCG ACCGCCAGCA GCAGATGATC GAGCGTGTGG AGAAGACCAC CGGGGACAAG
CGGACTCGCA TCCAGGGCCG TGTCGCCCAC CTCACTGGCA TCCATGCAGT GGAGGAAGTC
AGCCTGGAGG AGTTCTCAGC CCACCCATGT GCCCGTGACA ATGGTGGCTG CTCCCACATC
TGTATTGCCA AGGGTGATGG GACACCACGG TGCTCATGCC CAGTCCACCT CGTGCTCCTG
CAGAACCTGC TGACCTGTGG AGAGCCGCCC ACCTGCTCCC CGGACCAGTT TGCATGTGCC
ACAGGGGAGA TCGACTGTAT CCCCGGGGCC TGGCGCTGTG ACGGCTTTCC CGAGTGCGAT
GACCAGAGCG ACGAGGAGGG CTGCCCCGTG TGCTCCGCCG CCCAGTTCCC CTGCGCGCGG
GGTCAGTGTG TGGACCTGCG CCTGCGCTGC GACGGCGAGG CAGACTGTCA GGACCGCTCA
GACGAGGCGG ACTGTGACGC CATCTGCCTG CCCAACCAGT TCCGGTGTGC GAGCGGCCAG
TGTGTCCTCA TCAAACAGCA GTGCGACTCC TTCCCCGACT GTATCGACGG CTCCGACGAG
CTCATGTGTG AAATCACCAA GCCGCCCTCA GACGACAGCC CGGCCCACAG CAGTGCCATC

43/89

FIG. 24D

HBMGI_2AS

GGGCCC GTCA TTGGCATCAT CCTCTCTCTC TTCGTCATGG GTGGTGTCTA TTTTGTGTGC
CAGCGCGTGG TGTGCCAGCG CTATGCGGGG GCCAACGGGC CCTTCCCGCA CGAGTATGTC
AGCGGGACCC CGCACGTGCC CCTCAATTTC ATAGCCCCGG GCGGTTCCCA GCATGGCCCC
TTCACAGGCA TCGCATGCGG AAAGTCCATG ATGAGCTCCG TGAGCCTGAT GGGGGGCCGG
GGCGGGGTGC CCCTCTACGA CCGGAACCAC GTCACAGGGG CCTCGTCCAG CAGCTCGTCC
AGCACGAAGG CCACGCTGTA CCCGCCGATC CTGAACCCGC CGCCCTCCCC GGCCACGGAC
CCCTCCCTGT ACAACATGGA CATGTTCTAC TCTTCAAACA TTCCGGCCAC TGCGAGACCG
TACAGGCCCT ACATCATTCG AGGAATGGCG CCCCCGACGA CGCCCTGCAG CACCGACGTG
TGTGACAGCG ACTACAGCGC CAGCCGCTGG AAGGCCAGCA AGTACTACCT GGATTTGAAC
TCGGACTCAG ACCCCTATCC ACCCCCACCC ACGCCCCACA GCCAGTACCT GTCGGCGGAG
GACAGCTGCC CGCCCTCGCC CGCCACCGAG AGGAGCTACT TCCATCTCTT CCCGCCCCCT
CCGTCCCCCT GCACGGACTC ATCCTGACCT CGGCCGGGCC ACTCTGGCTT CTCTGTGCCC
CTGTAAATAG TTTTAAATAT GAACAAAGAA AAAAATATAT TTTATGATTT AAAAAATAAA
TATAATTGGG ATTTTAAAAA CATGAGAAAT GTGAACTGTG ATGGGGTGGG CAGGGCTGGG
AGAACTTTGT ACAGTGGAGA AATATTTATA AACTTAATTT TGTA AACAG AAAAAAAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA GCGGCCGC

44/89

FIG. 25A

ZMAXGI_3AS

TCTAGACTCG AGGCGGCCGC CCATTGTGCA CTAAAGCGCC GCCGCCGCGC CATGGAGCCC
GAGTGAGCGC GGC GCGGGCC CGTCCGGCCG CCGGACAACA TGGAGGCAGC GCCGCCCGGG
CCGCCGTGGC CGCTGCTGCT GCTGCTGCTG CTGCTGCTGG CGCTGTGCGG CTGCCCCGCC
CCCGCCGCGG CCTCGCCGCT CCTGCTATTT GCCAACCGCC GGGACGTACG GCTGGTGGAC
GCCGGCGGAG TCAAGCTGGA GTCCACCATC GTGGTCAGCG GCCTGGAGGA TGGCGCCGCA
GTGGACTTCC AGTTTTCCAA GGGAGCCGTG TACTGGACAG ACGTGAGCGA GGAGGCCATC
AAGCAGACCT ACCTGAACCA GACGGGGGCC GCCGTGCAGA ACGTGGTCAT CTCCGGCCTG
GTCTCTCCCG ACGGCCTCGC CTGCGACTGG GTGGGCAAGA AGCTGTACTG GACGGACTCA
GAGACCAACC GCATCGAGGT GGCCAACCTC AATGGCACAT CCCGGAAGGT GCTCTTCTGG
CAGGACCTTG ACCAGCCGAG GGCCATCGCC TTGGACCCCG CTCACGGGTA CATGTACTGG
ACAGACTGGG GTGAGACGCC CCGGATTGAG CGGGCAGGGA TGGATGGCAG CACCCGGAAG
ATCATTGTGG ACTCGGACAT TTA CTGGCCC AATGGACTGA CCATCGACCT GGAGGAGCAG
AAGCTCTACT GGGCTGACGC CAAGCTCAGC TTCATCCACC GTGCCAACCT GGACGGCTCG
TTCCGGCAGA AGGTGGTGGA GGGCAGCCTG ACGCACCCCT TCGCCCTGAC GCTCTCCGGG
GACACTCTGT ACTGGACAGA CTGGCAGACC CGCTCCATCC ATGCCTGCAA CAAGCGCACT
GGGGGGAAGA GGAAGGAGAT CCTGAGTGCC CTCTACTCAC CCATGGACAT CCAGGTGCTG
AGCCAGGAGC GGCAGCCTTT CTTCCACACT CGCTGTGAGG AGGACAATGG CGGCTGCTCC
CACCTGTGCC TGCTGTCCCC AAGCGAGCCT TTCTACACAT GCGCCTGCCC CACGGGTGTG
CAGCTGCAGG ACAACGGCAG GACGTGTAAG GCAGGAGCCG AGGAGGTGCT GCTGCTGGCC
CGGCGGACGG ACCTACGGAG GATCTCGCTG GACACGCCGG ACTTCACCGA CATCGTGCTG
CAGGTGGACG ACATCCGGCA CGCCATTGCC ATCGACTACG ACCCGCTAGA GGGCTATGTC
TACTGGACAG ATGACGAGGT GCGGGCCATC CGCAGGGCGT ACCTGGACGG GTCTGGGGCG
CAGACGCTGG TCAACACCGA GATCAACGAC CCCGATGGCA TCGCGGTCGA CTGGGTGGCC
CGAAACCTCT ACTGGACCGA CACGGGCACG GACCGCATCG AGGTGACGCG CCTCAACGGC

45/89

FIG. 25B

ZMAXGI_3AS

ACCTCCCGCA AGATCCTGGT GTCGGAGGAC CTGGACGAGC CCCGAGCCAT CGCACTGCAC
CCCGTGATGG GCCTCATGTA CTGGACAGAC TGGGGAGAGA ACCCTAAAAT CGAGTGTGCC
AACTTGATG GGCAGGAGCG GCGTGTGCTG GTCAATGCCT CCCTCGGGTG GCCCAACGGC
CTGGCCCTGG ACCTGCAGGA GGGGAAGCTC TACTGGGGAG ACGCCAAGAC AGACAAGATC
GAGGTGATCA ATGTTGATGG GACGAAGAGG CGGACCCTCC TGGAGGACAA GCTCCCGCAC
ATTTTCGGGT TCACGCTGCT GGGGGACTTC ATCTACTGGA CTGACTGGCA GCGCCGCAGC
ATCGAGCGGG TGCACAAGGT CAAGGCCAGC CGGGACGTCA TCATTGACCA GCTGCCCCGAC
CTGATGGGGC TCAAAGCTGT GAATGTGGCC AAGGTCGTCTG GAACCAACCC GTGTGCGGAC
AGGAACGGGG GGTGCAGCCA CCTGTGCTTC TTCACACCCC ACGCAACCCG GTGTGGCTGC
CCCATCGGCC TGGAGCTGCT GAGTGACATG AAGACCTGCA TCGTGCCTGA GGCCTTCTTG
GTCTTCACCA GCAGAGCCGC CATCCACAGG ATCTCCCTCG AGACCAATAA CAACGACGTG
GCCATCCCGC TCACGGGCGT CAAGGAGGCC TCAGCCCTGG ACTTTGATGT GTCCAACAAC
CACATCTACT GGACAGACGT CAGCCTGAAG ACCATCAGCC GCGCCTTCAT GAACGGGAGC
TCGGTGGAGC ACGTGGTGGA GTTTGGCCTT GACTACCCCG AGGGCATGGC CGTTGACTGG
ATGGGCAAGA ACCTCTACTG GGCCGACACT GGGACCAACA GAATCGAAGT GGC GCGGCTG
GACGGGCAGT TCCGGCAAGT CCTCGTGTGG AGGGACTTGG ACAACCCGAG GTCGCTGGCC
CTGGATCCCA CCAAGGGCTA CATCTACTGG ACCGAGTGGG GCGGCAAGCC GAGGATCGTG
CGGGCCTTCA TGGACGGGAC CAACTGCATG ACGCTGGTGG ACAAGGTGGG CCGGGCCAAC
GACCTCACCA TTGACTACGC TGACCAGCGC CTCTACTGGA CCGACCTGGA CACCAACATG
ATCGAGTCGT CCAACATGCT GGGTCAGGAG CGGGTCGTGA TTGCCGACGA TCTCCCGCAC
CCGTTTCGGTC TGACGCAGTA CAGCGATTAT ATCTACTGGA CAGACTGGAA TCTGCACAGC
ATTGAGCGGG CCGACAAGAC TAGCGGCCGG AACC GCACCC TCATCCAGGG CCACCTGGAC
TTCGTGATGG ACATCCTGGT GTTCCACTCC TCCCGCCAGG ATGGCCTCAA TGA CTGTATG

46/89

FIG. 25C

ZMAXGI_3AS

CACAACAACG GGCAGTGTGG GCAGCTGTGC CTTGCCATCC CCGGCGGCCA CCGCTGCGGC
TGCGCCTCAC ACTACACCCT GGACCCAGC AGCCGCAACT GCAGCCCGCC CACCACCTTC
TTGCTGTTCA GCCAGAAATC TGCCATCAGT CGGATGATCC CGGACGACCA GCACAGCCCC
GATCTCATCC TGCCCCTGCA TGGACTGAGG AACGTCAAAG CCATCGACTA TGACCCACTG
GACAAGTTCA TCTACTGGGT GGATGGGCGC CAGAACATCA AGCGAGCCAA GGACGACGGG
ACCCAGCCCT TTGTTTTGAC CTCTCTGAGC CAAGGCCAAA ACCCAGACAG GCAGCCCCAC
GACCTCAGCA TCGACATCTA CAGCCGGACA CTGTTCTGGA CGTGCGAGGC CACCAATACC
ATCAACGTCC ACAGGCTGAG CGGGGAAGCC ATGGGGGTGG TGCTGCGTGG GGACCGCGAC
AAGCCCAGGG CCATCGTCGT CAACGCGGAG CGAGGGTACC TGTACTTCAC CAACATGCAG
GACCGGGCAG CCAAGATCGA ACGCGCAGCC CTGGACGGCA CCGAGCGCGA GGTCTCTTTC
ACCACCGGCC TCATCCGCCC TGTGGCCCTG GTGGTAGACA ACACACTGGG CAAGCTGTTC
TGGGTGGACG CGGACCTGAA GCGCATTGAG AGCTGTGACC TGTCAGGGGC CAACCGCCTG
ACCCTGGAGG ACGCCAACAT CGTGCGAGCT CTGGGCCTGA CCATCCTTGG CAAGCATCTC
TACTGGATCG ACCGCCAGCA GCAGATGATC GAGCGTGTGG AGAAGACCAC CGGGGACAAG
CGGACTCGCA TCCAGGGCCG TGTCGCCAC CTTACTGGCA TCCATGCAGT GGAGGAAGTC
AGCCTGGAGG AGTTCTCAGC CCACCCATGT GCCCGTGACA ATGGTGGCTG CTCCCACATC
TGTATTGCCA AGGGTGATGG GACACCACGG TGCTCATGCC CAGTCCACCT CGTGCTCCTG
CAGAACCTGC TGACCTGTGG AGAGCCGCC ACCTGCTCCC CGGACCAGTT TGCATGTGCC
ACAGGGGAGA TCGACTGTAT CCCCGGGGCC TGGCGCTGTG ACGGCTTTCC CGAGTGCGAT
GACCAGAGCG ACGAGGAGGG CTGCCCCGTG TGCTCCGCCG CCAGTTCCCC TGC GCGCGG
GGTCAGTGTGT GGACCTGCGC CTGCGCTGCG ACGGCGAGGC AGACTGTCAG GACCGCTCA
GACGAGGCGGA CTGTGACGCC ATCTGCCTGC CCAACCAGTT CCGGTGTGCG AGCGGCCAG
TGTGTCCTCAT CAAACAGCAG TCGACTCCT TCCCCGACTG TATCGACGGC TCCGACGAG
CTCATGTGTGA AATCACCAAG CCGCCCTCAG ACGACAGCCC GGCCCACAGC AGTGCCATC

47/89

FIG. 25D

ZMAXGI_3AS

GGGCCCCGTCA TTGGCATCAT CCTCTCTCTC TTCGTCATGG GTGGTGTCTA TTTTGTGTGC
CAGCGCGTGG TGTGCCAGCG CTATGCGGGG GCCAACGGGC CCTTCCCGCA CGAGTATGTC
AGCGGGACCC CGCACGTGCC CCTCAATTTT ATAGCCCCGG GCGGTTCCCA GCATGGCCCC
TTCACAGGCA TCGCATGCGG AAAGTCCATG ATGAGCTCCG TGAGCCTGAT GGGGGGCCGG
GGCGGGGTGC CCCTCTACGA CCGGAACCAC GTCACAGGGG CCTCGTCCAG CAGCTCGTCC
AGCACGAAGG CCACGCTGTA CCCGCCGATC CTGAACCCGC CGCCCTCCCC GGCCACGGAC
CCCTCCCTGT ACAACATGGA CATGTTCTAC TCTTCAAACA TTCCGGCCAC TGCAGACCG
TACAGGCCCT ACATCATTCG AGGAATGGCG CCCCCGACGA CGCCCTGCAG CACCGACGTG
TGTGACAGCG ACTACAGCGC CAGCCGCTGG AAGGCCAGCA AGTACTACCT GGATTTGAAC
TCGGACTCAG ACCCCTATCC ACCCCCACCC ACGCCCCACA GCCAGTACCT GTCGGCGGAG
GACAGCTGCC CGCCCTCGCC CGCCACCGAG AGGAGCTACT TCCATCTCTT CCCGCCCCCT
CCGTCCCCCT GCACGGACTC ATCCTGACCT CGGCCGGGCC ACTCTGGCTT CTCTGTGCCC
CTGTAAATAG TTTTAAATAT GAACAAAGAA AAAAATATAT TTTATGATTT AAAAAATAAA
TATAATTGGG ATTTTAAAAA CATGAGAAAT GTGAACTGTG ATGGGGTGGG CAGGGCTGGG
AGAACTTTGT ACAGTGGAGA AATATTTATA AACTTAATTT TGTA AACAG AAAAAAAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA GCGGCCGC

FIG. 26A

Amino-Acid Sequences.

1	MEA...APPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLV	46
1	METAPTRAPPPPPPPPLLLLV.....LYCSLVPAAASPLLLFANRRDVRLV	45
47	DAGGVKLESTIVVSGLEDAAAVDFQFSKGAVYWTDVSEEAIKQTYLNQTG	96
46	DAGGVKLESTIVASGLEDAAAVDFQFSKGAVYWTDVSEEAIKQTYLNQTG	95
97	AAVQNVVISGLVSPDGLACDWVGKKLYWTDSETNRIEVANLNGTSRKVL	146
96	AAAQNIIVISGLVSPDGLACDWVGKKLYWTDSETNRIEVANLNGTSRKVL	145
147	WQDL DQPRAI ALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIVDSDIYW	196
146	WQDL DQPRAI ALDPAHGYMYWTDWGEAPRIERAGMDGSTRKIIVDSDIYW	195
197	PNGLTIDLEE QKLYWADAKLSFIHRANLDG SFRQKVVEGSLTHPFALTLS	246
196	PNGLTIDLEE QKLYWADAKLSFIHRANLDG SFRQKVVEGSLTHPFALTLS	245
247	GDTLYWTDWQTRSIHACNKRTGGKRKEILSALYSPMDIQVLSQERQPPFH	296
246	GDTLYWTDWQTRSIHACNKWTGEQRKEILSALYSPMDIQVLSQERQPPFH	295
297	TRCEEDNGGC SHLCLLSPSEPFYTCACPTGVQLQDNGRTCKAGAEVLLL	346
296	TPCEEDNGGC SHLCLLSPREPFYSCACPTGVQLQDNGKTCKTGAEVLLL	345
347	ARRTDLRRISL DTPDFTDIVLQVDDIRHAI AIDYDPLEGYVYWTDDEVRA	396
346	ARRTDLRRISL DTPDFTDIVLQVGDIRHAI AIDYDPLEGYVYWTDDEVRA	395
397	IRRAYLDGSGAQTLVNTEINDPDGIAVDWVARNLYWTDGTDRIEVTRLN	446
396	IRRAYLDGSGAQTLVNTEINDPDGIAVDWVARNLYWTDGTDRIEVTRLN	445
447	GTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGENPKIECANLDGQERRV	496
446	GTSRKILVSEDLDEPRAIVLHPVMGLMYWTDWGENPKIECANLDGRDRHV	495
497	LVNASLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTKRRTLLEDKLP	546
496	LVNTSLGWPNGLALDLQEGKLYWGDAKTDKIEVINIDGTKRKTLLLEDKLP	545
547	HIFGFTLLGDFIYWTDWQRRS IERVHKVKASRDVIIDQLPDL MGLKAVNV	596
546	HIFGFTLLGDFIYWTDWQRRS IERVHKVKASRDVIIDQLPDL MGLKAVNV	595
597	AKVVGTPNCADRNGGCSHL CFFTPHATRCGCPIGLELLSDMKTCIVPEAF	646
596	AKVVGTPNCADGNGGCSHL CFFTPRATKCGCPIGLELLSDMKTCI IPEAF	645
647	LVFTSRAAIHRISLETNNNDVAIPLTG VKEASALDFDVSNNH IYWTDVSL	696
646	LVFTSRATIIHRISLETNNNDVAIPLTG VKEASALDFDVSNNH IYWTDVSL	695

49/89

FIG. 26B

Alignment of Human and Mouse Zmax1 (LRP5)
Amino-Acid Sequences.

```

697 KTISRAFMNGSSVEHVVEFGLDYPEGMAVDWMGKNLYWADTGTNRIEVAR 746
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 KTISRAFMNGSSVEHVIEFGLDYPEGMAVDWMGKNLYWADTGTNRIEVAR 745
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
747 LDGQFRQVLVWRDLNPRSLALDPTKGYIYWTEWGGKPRIVRAFMGTNC 796
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
746 LDGQFRQVLVWRDLNPRSLALDPTKGYIYWTEWGGKPRIVRAFMGTNC 795
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
797 MTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLGQERVVIADDLP 846
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
796 MTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLGQERMVIADDLP 845
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
847 HPFGLTQYSDYIYWTDWNLHSIERADKTSGRNRTLQGHLD FVMDILVFH 896
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
846 YPFGLTQYSDYIYWTDWNLHSIERADKTSGRNRTLQGHLD FVMDILVFH 895
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
897 SSRQDGLNDÇMHNNQCGQLCLAIPGGHRCGCASHYTLDPSSRNCSPTT 946
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
896 SSRQDGLNDCVHSNGQCGQLCLAIPGGHRCGCASHYTLDPSSRNCSPPST 945
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
947 FLLFSQKSAISRMIPDDQHSPDLILPLHGLRNVKAIDYDPLDKFIYWVDG 996
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
946 FLLFSQKFAISRMIPDDQLSPDLVLPLHGLRNVKAINYDPLDKFIYWVDG 995
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
997 RQNIKRAKDDGTQPFVLTSLSQGQNPDRQPHDLSIDIYSRTLFWTCEATN 1046
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
996 RQNIKRAKDDGTQPSMLTSPSLSQSPDRQPHDLSIDIYSRTLFWTCEATN 1045
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1047 TINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGYLYFTNMQDRAAKIERA 1096
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1046 TINVHRLDGDAMGVVLRGDRDKPRAIVVNAERGYMYFTNMQDHAAKIERA 1095
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1097 ALDGTEREVLFTTGLIRPVALVVDNTLGKLFWVDADLKRIESCDLSGANR 1146
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1096 SLDGTEREVLFTTGLIRPVALVVDNALGKLFWVDADLKRIESCDLSGANR 1145
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1147 LTLEDANIVQPLGLTILGKHL YWIDRQQQMIERVEKTTGDKRTRIQGRVA 1196
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1146 LTLEDANIVQPVGLTVLGRHLYWIDRQQQMIERVEKTTGDKRTRVQGRVT 1195
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1197 HLTGIHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVL 1246
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1196 HLTGIHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVL 1245
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1247 LQNLLTCGEPTCSPDQFAÇATGEIDCIPGAWRCDFPEÇDDQSDEEGCP 1296
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1246 LQNLLTCGEPTCSPDQFACTTGEIDCIPGAWRCDFPEÇADQSDEEGCP 1295
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1297 VCSAAQFPCARGQCVDLRLRCDGEADCQDRSDEADCAIÇLPNQFRCASÇ 1346
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1296 VCSASQFPCARGQCVDLRLRCDGEADCQDRSDEANCDVCLPNQFRCTSG 1345
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1347 QCVLIKQQCDSFPDCIDGSDELMCEITKPPSDDSPAHSŞAIGPVIGIILŞ 1396
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1346 QCVLIKQQCDSFPDCADGSDELMCEINKPPSDDIPAHSŞAIGPVIGIILŞ 1395
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1397 LFVMGGVYFVCQRVVCQRYAGANGPFPHEYVSGTPHVPLNFIAPGGSQHG 1446
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1396 LFVMGGVYFVCQRVVCQRYTGASGFPFPHEYVGGAPHVPLNFIAPGGSQHG 1445
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

FIG. 26C

[illegible]

51/89

FIG. 27A

Alignment of Human LRP5 and LRP6

Top = LRP5, Bottom = LRP6

Underlined = region marked for antibody production

Percent Similarity: 83.542 Percent Identity: 70.901

```

1  MEAAPPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLVDAGG 50
1  .....MGAVLRSLLA.CSFCVLLRAAPLLLYANRRDLRLVDATN 38
51  VKLESTIVVSGLEDAAAVDFQFSKGAVYWTVDVSEEAIKQTYLNQTGAAVQ 100
39  GKENATIVVGGLEDAAAVDFVFSHGLIYWSVDVSEEAIKRTEFNKTE.SVQ 87
101  NVVISGLVSPDGLACDWVGKKLYWTDSETNRIEVANLNGTSRKVLFWQDL 150
88  NVVVSGLLSPDGLACDWLGEKLYWTDSETNRIEVSNDGSLRKVLFWQEL 137
151  DQPRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIVDSDIYWPNGL 200
138  DQPRAIALDPSSGFMWTDWGEVPKIERAGMDGSSRFIIINSEIYWPNGL 187
201  TIDLEEQKLYWADAKLSFIHRANLDGSFRQKVVEGSLTHPFALTLSGDTL 250
188  TLDYEEQKLYWADAKLNFHKSNDGTNRQAVVKGSLPHPFALTTFEDIL 237
251  YWTDWQTRS IHACNKRTGGKRKEILSALYS PMDIQVLSQERQPFHTRCE 300
238  YWTDWSTHSILACNKYTGEGLREIHS DIFSPMDIHAFSQQRQPNATNPCG 287
301  EDNGGC SHLCLLSPSEPFYTCACPTGVQLQDNGRTCKAGAEVLLLARRT 350
288  IDNGGC SHLCLMSPVKPFYQCACPTGVKLLENGKTCCKDGATELLLLARRT 337
351  DLRRISLDTPDFTDIVLQVDDIRHAI AIDYDPLEGYVYWTDDDEVRAIRRA 400
338  DLRRISLDTPDFTDIVLQLEDIRHAI AIDYDPVEGYIYWTDDDEVRAIRRS 387
401  YLDGSGAOTLVNTEINDPDGIAVDWVARNLYWTDGTGTDRIEVTRLNGTSR 450
388  FIDGSGSQFVVT AQIAHPDGI AVDWVARNLYWTDGTGTDRIEVTRLNGTMR 437
451  KILVSEDLDEPRAIALHPVMGLMYWTDWG ENPKIECANLDGQERRVLVNA 500
438  KILISEDLEEPRAIVLDPVMGYMYWTDWGEIPKIERAALDGS DRVVLVNT 487
501  SLGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTKRRTLLEDKLP HIFG 550
488  SLGWPNGLALDYDEGKIYWGDAKTDKIEVMNTDGTGRRVLVEDKIP HIFG 537
551  FTLLGDFIYWTWQRRS IERVHKVKASRDVIDQLPDL MGLKAVNVAKVV 600
538  FTLLGDYVYWTWQRRS IERVHKRSAEREVIDQLPDL MGLKATNVHRVI 587
601  GTNPCADRNGGCSHL CFFTPHATRCGCP IGLELLSDMKT CIVPEAFLVFT 650
588  GSNPCAEEENGCSHLCLYRPQGLRCACPIGFELISDMKT CIVPEAFLIFS 637

```

52/89

FIG. 27B

Alignment of Human LRP5 and LRP6

```

651 SRAAIHRISLETNNNDVAIPLTGVKEASALDFDVSNNHIYWTDVSLKTIŠ 700
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
638 RRADIRRISLETNNNNVAIPLTGVKEASALDFDVTDNRIYWTDISLKTIS 687
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
701 RAFMNGSSVĖHVVEFGLDYPEGMAVDWMGKNLYWADTGTNRIEVARLDGQ 750
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
688 RAFMNGSALEHVVEFGLDYPEGMAVDWLGNLYWADTGTNRIEVSKLDGQ 737
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
751 FRQVLVWRDLĎNPRSLALDPTKGYYWTEWGGKPRIVRAFMDGTNCMTLV 800
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
738 HRQVLVWKDLDSPRALALDPAEGFMYWTEWGGKPKIDRAAMDGSERTTLV 787
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
801 ĐKVGRANDLTĎIDYADQRLYWTDLDTNMIĖŠSNMLGQERVVIADDLPHPFĖ 850
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
788 PNVGRANGLTĎIDYAKRRLYWTDLDTNLIĖŠSNMLGLNREVIADDLPHPFĖ 837
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
851 LTQYSDYIYWTDWNLHSIERADKTSGRNRĎLIQGHLDĖMDILVFHSSRQ 900
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
838 LTQYQDYIYWTDWSRRSIERANKTSQQRĎTIIQGHLDYVMDILVFHSSRQ 887
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
901 DGLNDCMHNNĖGCGQLCLAIP.GGHRCGCASHYTLDPSSRNCSPPTTFLĎ 949
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
888 ŠGWNECASSNGHCSHLCLAVPVGGFVCGPCAHYSLNADNRĎTCSAPTFLĎ 937
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
950 FSQKSAISRMĎIPDDQHSPDLĎILPLHGLRNVKAIDYDPLDKĎFIYWVDGRQŎ 999
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
938 FSQKSAINRMVIDEQQSPDIĎILPIHSLRNVRAIDYDPLDKQLYWĎDSRQŎ 987
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1000 .IKRAKDDGTQPF.VLTSLSQGONPDRQPHDLŠIDIYSRĎTLFWTCEATŎT 1047
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
988 MIRKAQEDGŠQGFTVVVŠSVPSQNLEIQPYDLŠIDIYSRYIYWĎTCEATNV 1037
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1048 INVHRLSGĖAMGVVLRGDRĐKPRAIVVNAĖRGYLYFTNMĐDRAAKIERAĖ 1097
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1038 INVTRLĐGRŠGVVLKGEQDRPRAIVVNPEKGYMYFTNLQERSPKIERAĖ 1087
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1098 LDGTEREVLĎFTTGLIRPVALVVDNTLGKLĎFWVDADLKRIĖSCDLŠGANRL 1147
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1088 LDGTEREVLĎFFSGLŠKPIALALĐSRLGKLĎFWADŠDLRRIĖŠSDLŠGANRI 1137
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1148 TLEDANIVQPLGLTILGKHLYWĎDRQQQMĎIERVEKTTGDĐKRTRIQGRVAĎ 1197
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1138 VLEDŠNILQPVGLTVFENWLYWĎDKQQQMĎIEKIDMTGREGRTKVQARIAQ 1187
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1198 LTGIHAVEĖVŠLEEFSAHPCĎARDNGGCSHĎCIAKGDGTPRĎCSCPVLVLĎ 1247
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1188 LSDIHAVKELNLQEYRQHPCAQĐNGGCSHĎICLVKGDGTTRĎCSCPMHLVLĎ 1237
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1248 QNLLTCGEPPĎTCSPĐQFACĎTGEIDCIPGAWRCDGFPECDĐQSDEEGCPV 1297
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1238 QDELSCGEPPĎTCSPQQFTCFTGEIDCIPVAWRCDGFTECEDHŠDELNCPV 1287
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1298 CSĖAAQFPCĖRGOCVDLRLRĎCDGEADCQRŠDEADCDAICĎLPNQFRCASGQ 1347
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1288 CSĖŠQFQCASGQCIDGALRCNGDANCQDKŠDEKNCEVLCLIDQFRCANGQ 1337
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1348 CVLIKQOCĐŠFPDCIDGŠDELMCĖITKPPŠDDSPAHSŠAIGPVIGIILŠ 1397
    |||||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1338 CIGKHKKCDHNVDČSDKŠDELDCYPTĖĖP...APQATNTVGSVIGVIVTI 1384
    .|||.||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

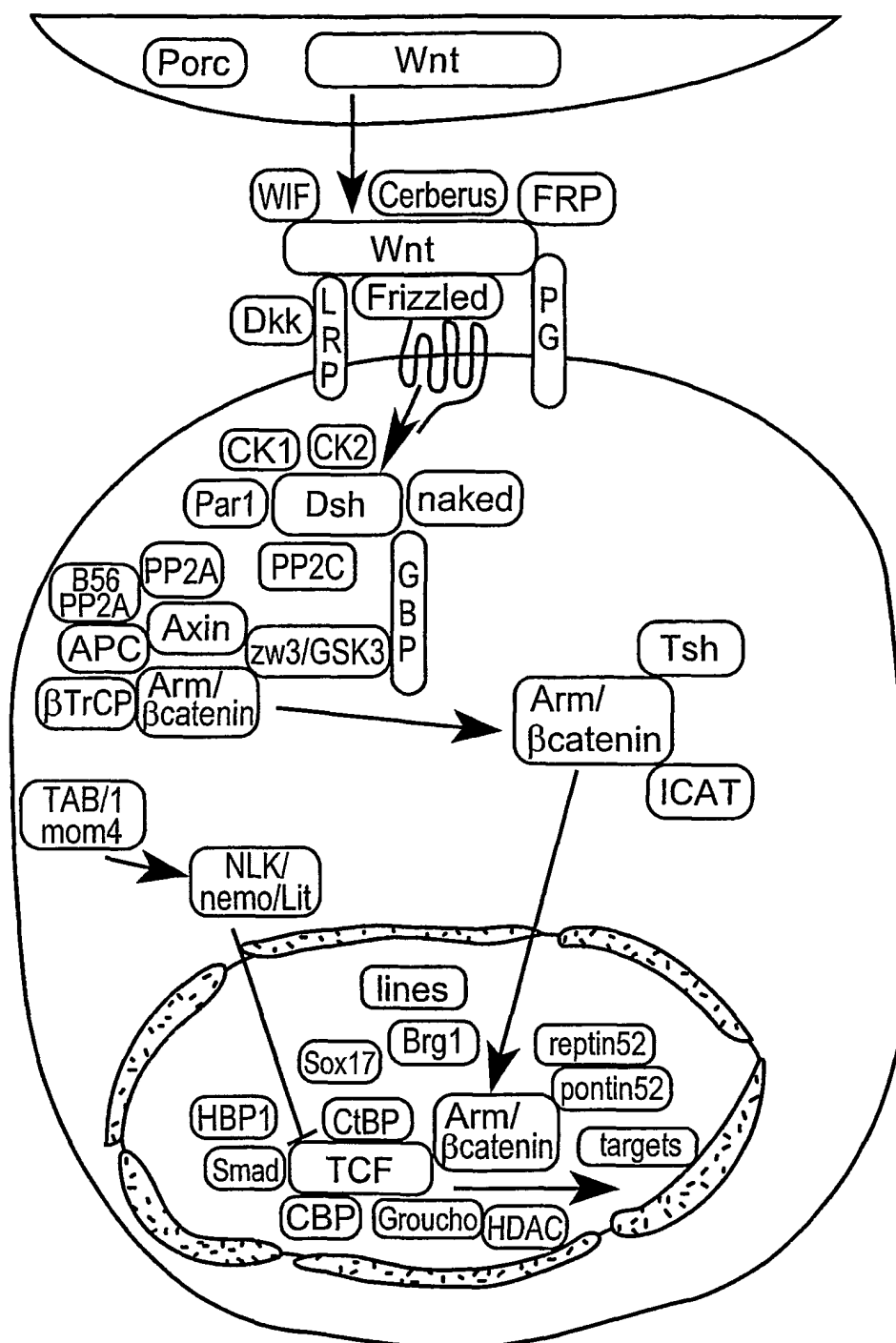

FIG. 27C

```

1398 FVMGGVYFVCQRVVCQRYAGANGPFPHEYVS.GTPHVPLNFIAPGGSQHĠ 1446
    |||.|||:||||:|.|||.:.:.:.:.|||.:.:.:.:.|||.
1385 FVSGTVYFICQRMCLCPRMKGDGETMTNDYVVHGPASVPLGYVPHPSLSG 1434
1447 PFTGIACGKŠMSSSVSLMGGRGGVPLYDRNHVTGASSSSŠSSTKATLYPP 1496
    .:.|.:.|.||||:||||:|.||||:|.|.|||||.|||||.|||||.|||||.:.:.|.:.|
1435 SLPGMSRGKSMISSLSIMGGSSGPP.YDRAHVTGASSSSSSSTKGTYFPA 1483
1497 ILNPPSPATDPSLYNMDMFYSSNIPATAR.PYRPYIIRGMAPPTTPCŠ 1544
    |||.|.|.|.:.|.|.:.|.|||||.|.|.|.|||||.|.|.|.|||||.|.|.|.
1484 ILNPPSPATERSHYTMEFGYSSNSPŠTHRSYSYRPYSYRHFAPPTTPCŠ 1533
1545 TDVCSDYSASRWKAS....KYYLDLNSĐSDPYPPPPTPHSQYLSAE.. 1587
    |||.|.|.|.:.|.|.|.|.|||||.|.|.|.|||||.|.|.|.|||||.|.|.|.
1534 TDVCSDSYAPŠRRMTSVATAKGYTSDLNVDSEPVPPPPTPRSQYLSAEEN 1583
1588 .DSCPPSPATERSY.FHLFPPPPSPCTDŠŠ 1615
    :.|||||.|.|||||.|||.|||||.|||||.|||||.
1584 YESCPPSPYTERSYSHHLYPPPPSPCTDŠŠ 1613

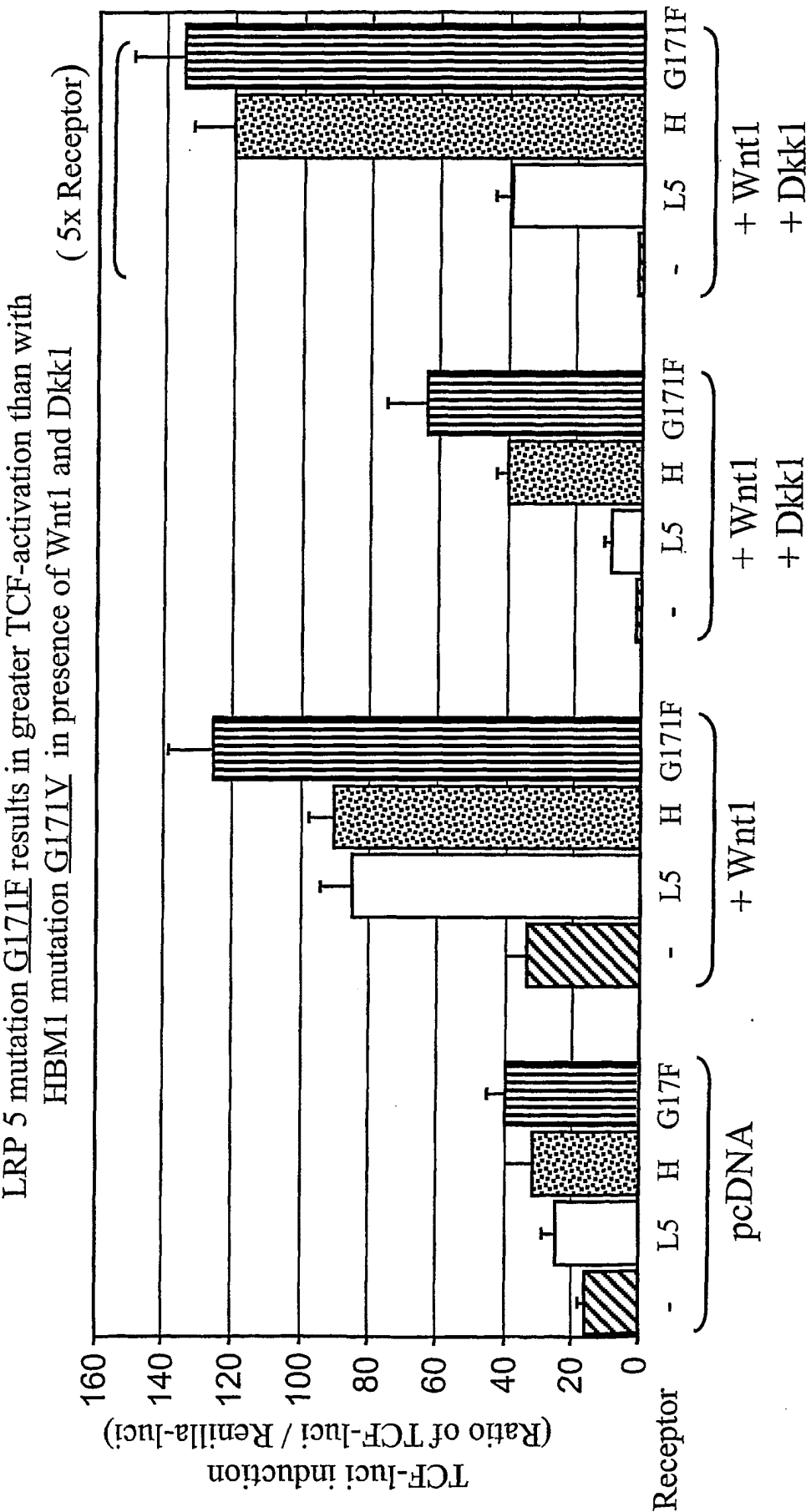
```

54/89



Model of Wnt signaling

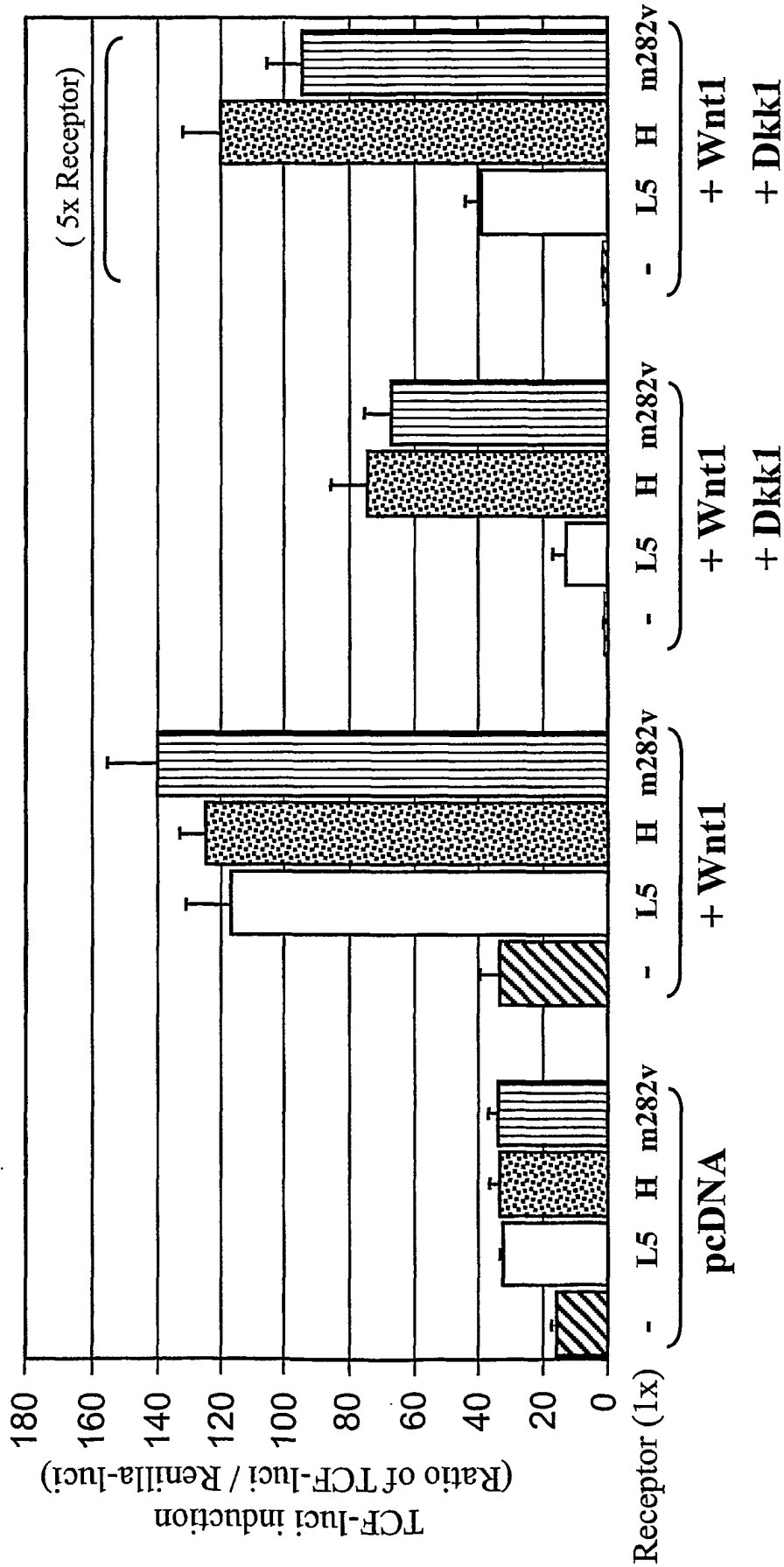
FIG. 28



• G171F mutation involves the ringed R group (F) alteration and leads to marginally greater TCF-luciferase activation than that with HBM1 mutation G171V.

FIG. 29

LRP5-blade1 Mutant M282V leads to HBM1 type TCF-signal
with Wnt1 and Dkk1 in U2-OS cells



- In blade 1, propeller 1, M282 is at the accessible interior position.
- It is conserved in propellers 1-3

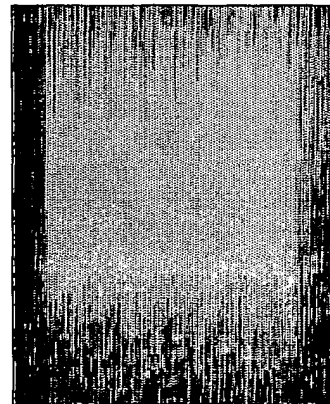
FIG. 30

Gene	Genbank Accession #	Protein Accession #
granulin	M75161	AAA58617
similar to cys/His rich protein	BC004544	AAH04544
IGF-BINDING PROTEIN 6	M69054	AAA88070
latent TGFb binding protein 4	AF051344	AAC39879
NOTCH 2	AF315356	AAG37073
fibulin 1	X53743	CAA37772
MDC15 (ADAM15)	U46005	AAC51112
DKFZp761G02121(notch1 Ca++ binding like)	AL137311	CAB70690
chordin	AF076612	AAC69835
fibronectin 1	U42594	AAD00019
MG50(melanoma associated antigen)	AF200348	AAF06354
unknown (notch 4-like)	AX068260	CAC27245
Slit 1	AB017167	BAA35184
tomoregulin (agarin repeat homology)	AB004064	BAA90820
sprouty 1	AF041037	AAC39566
sprouty 2	AF039843	AAC04258
NOV1	X96584	CAA65403
agrin	AF016903	AAC39776
fibrillin 1	L13923	AAB02036
thrombospondin1	X04665	CAA28370
ADAM19	AF134707	AAF22162
Nafl alpha	AJ011895	CAA09855
laminin alpha 5	Z95636	CAB09137
CRIM1	AF167706	AAF34409
nidogen	M30269	AAA59932
fibulin-2	X82494	CAA57876
thrombospondin 2	L12350	AAA03703
KIAA1323	AB037744	BAA92561
fibrillin-2	U03272	AAA18950
MEGF9	AB011542	BAA32470
integrin beta 1	X07979	CAA30790
matrilin-2 precursor	U69263	AAC51260
tenascin	X56160	A32160

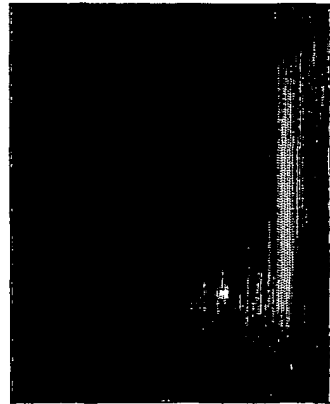
FIG. 31

58/89

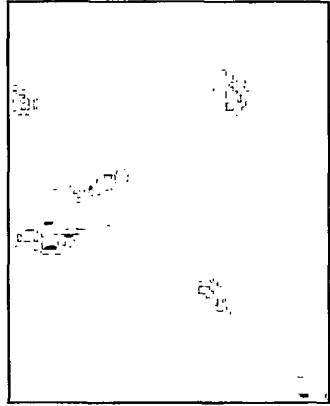
FIG. 32
Affinity Purified 69546/47



Non-infected



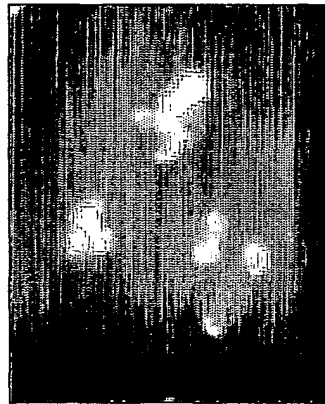
LRP5 virus infected,
fluorescence



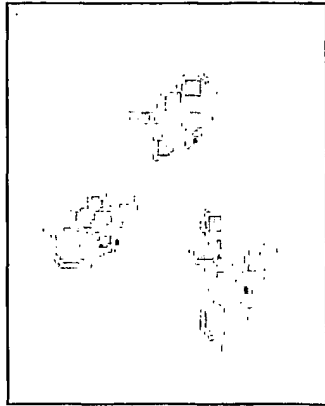
LRP5 virus infected, phase



Non-infected



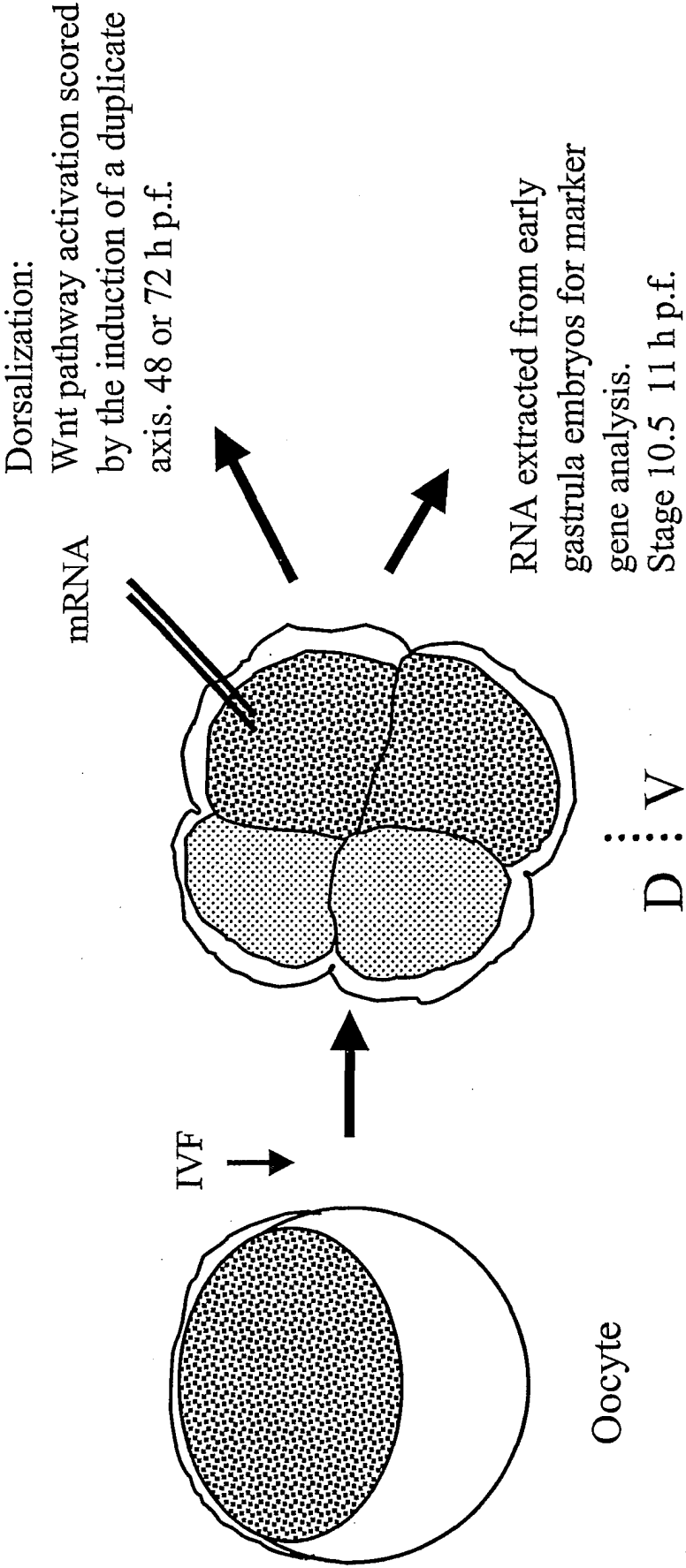
HBM virus infected,
fluorescence



HBM virus infected, phase

Antibody to: aa 165-177 (Mutation)

Xenopus Embryo Assay for Wnt Activity



RNA Injection into
ventral blastomere
Stage 4 2 h p.f.

FIG. 33

In the Presence of Wnt5a, HBM1 is More Potent than
Zmax as a Stimulator of Wnt Signaling

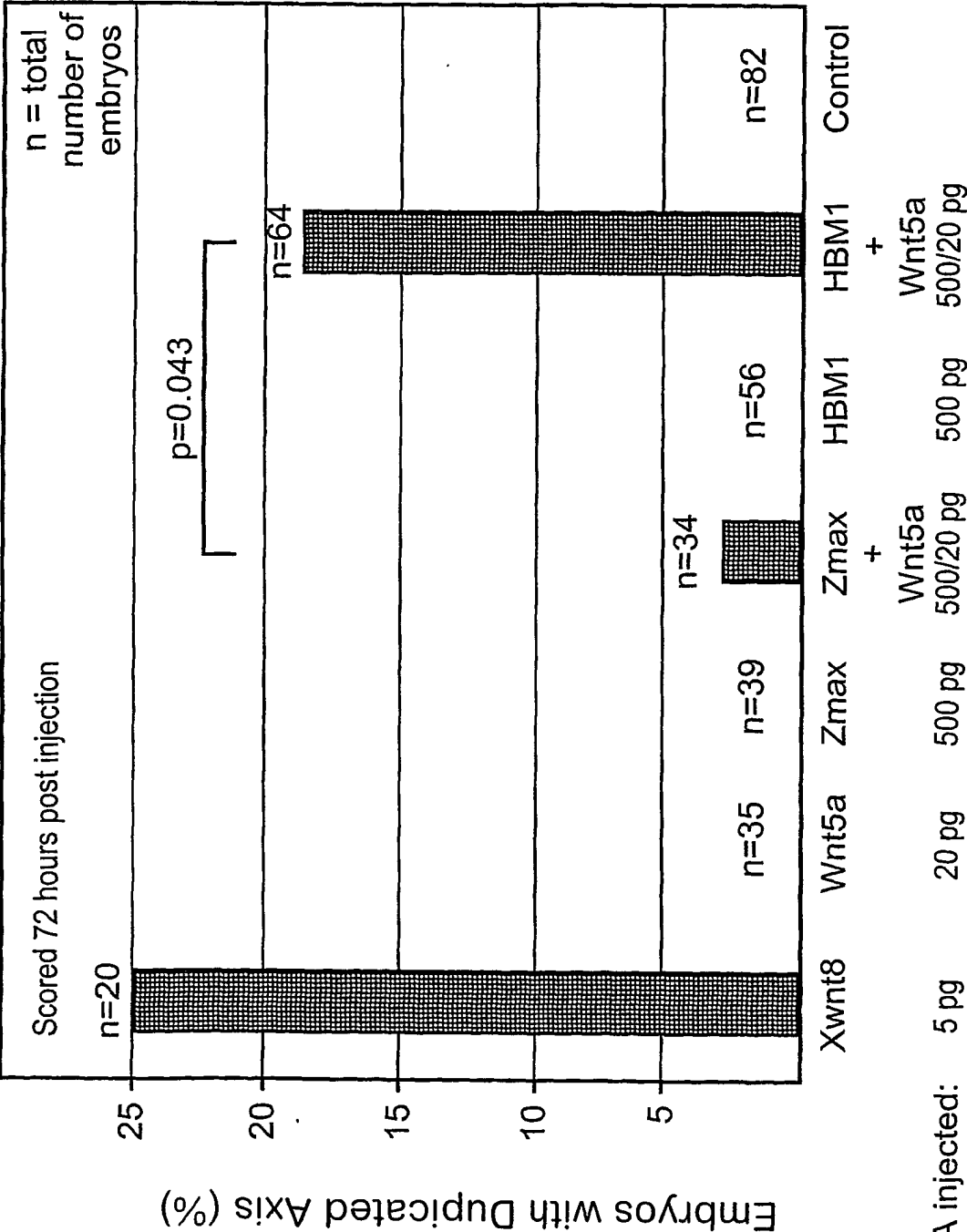


FIG. 34

61/89

Both Zmax and HBM1, in the presence of Wnt5a, induce secondary axis formation in Xenopus (photos at 48 hrs post-injection)

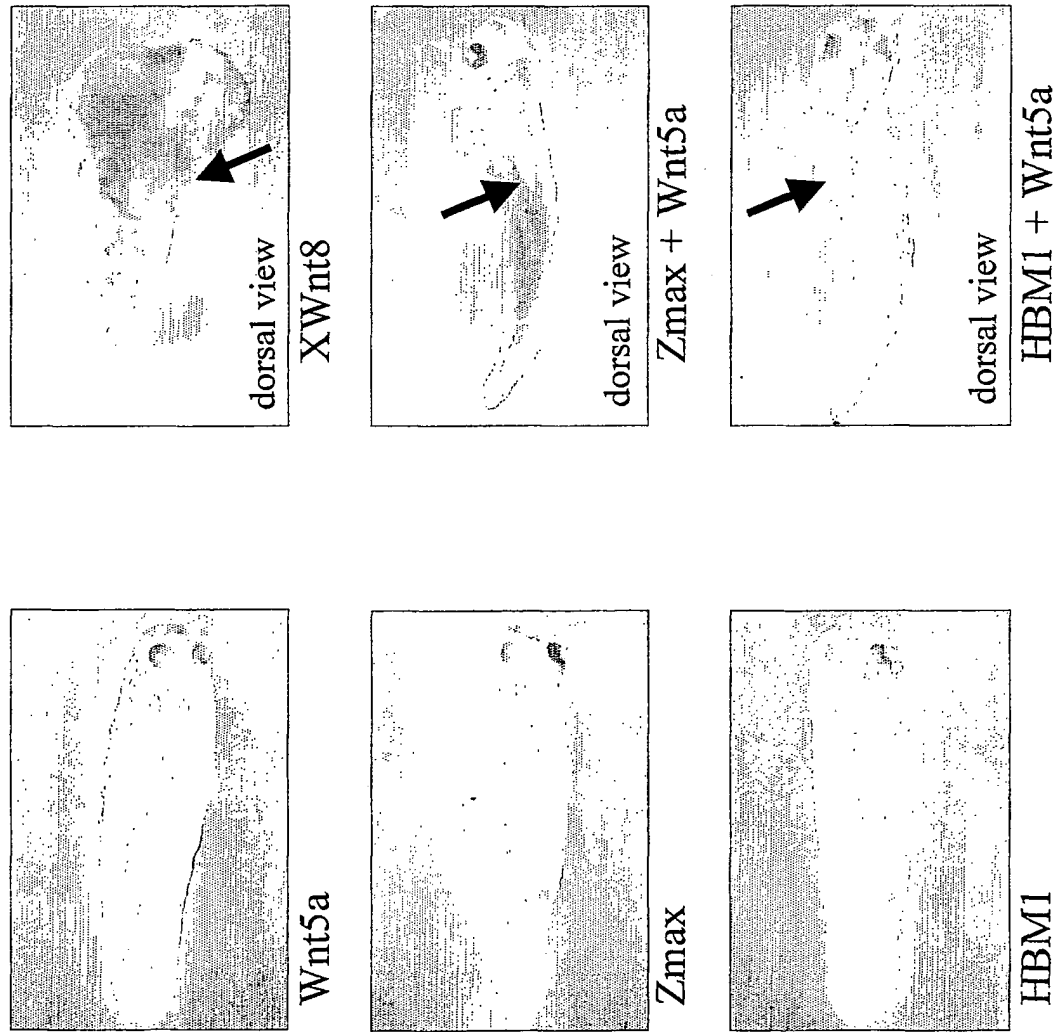


FIG. 35

Human Dkk-1 Represses the Canonical Wnt Pathway

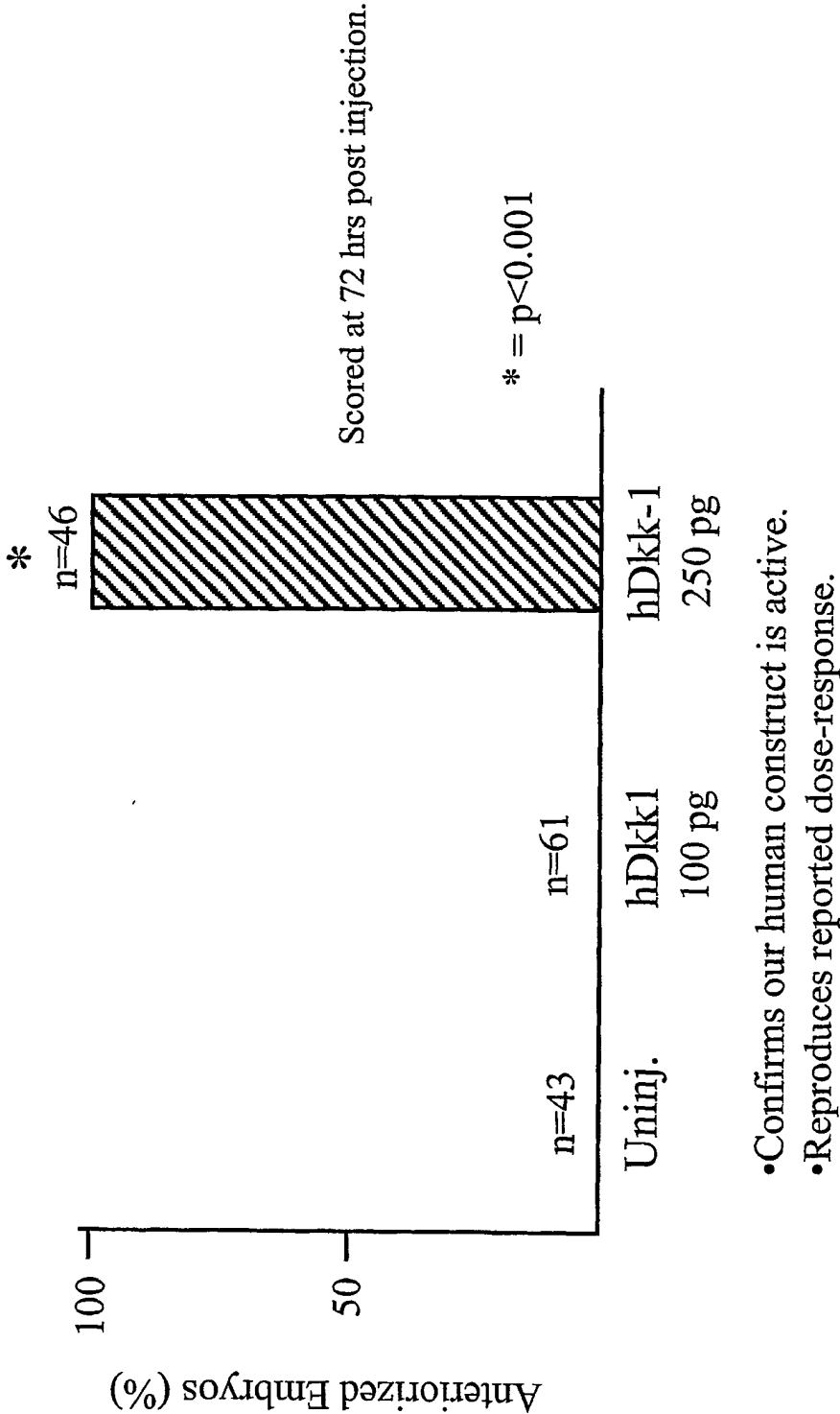


FIG. 36

63/89

hDkk1 Represses Zmax- but Not HBM1-Mediated Wnt Signaling

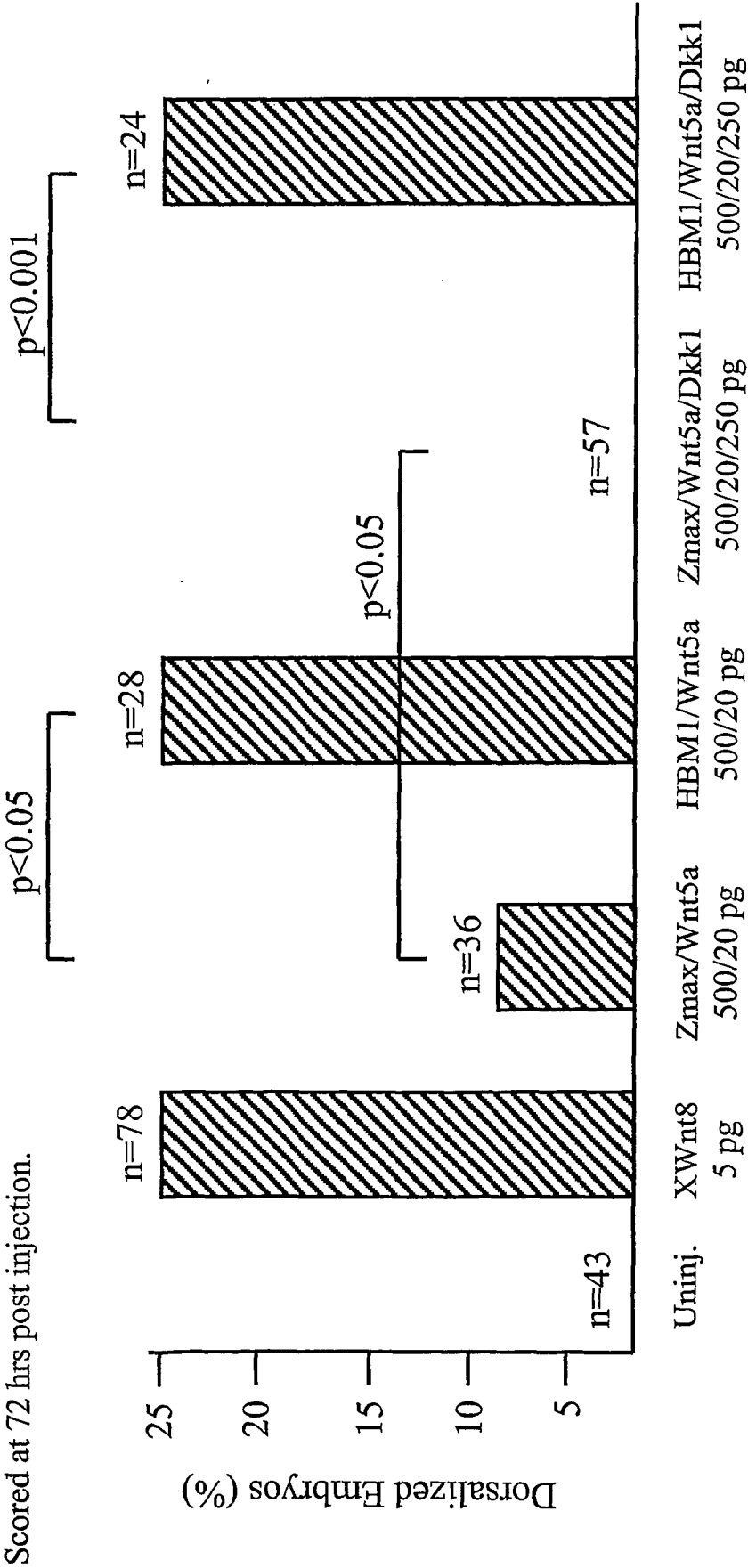


FIG. 37

64/89

zmaxLBD
1,4
screen X
peptide
library

Peptide Number	Peptide Sequence	No. of Hits
9	VVLCSRCGRLWRWSCG	1
12	EVRQVTCIRCRRGFL	1
13	GGGGMWEAWSCYACG	1
14	GWRWCGRCGALWWRRV	3

FIG. 38

65/89

Listed are the pcDNA3.1 construct names followed by the DNA sequence

OST258 (control for OST 259-OST262 and OST264, OST265)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCTACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA

OST259

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCTACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA
CTGACGACAGTTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTC
GATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGTGGTTCTGTGTTTCGCGTTG
TGGGCGTTTGTGGCGGTGGTCGTGTGGGACTAGTGGTCCGTGCAAAATGATCG
CCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCA
AAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGG
GTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA
GCGGCCGC

OST260

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCTACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA
CTGACGACAGTTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTC
GATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGGGTGGCGGTGGTGTGGTCG
GTGTGGGGCTTTGTGGTGGCGGCGTGTACTAGTGGTCCGTGCAAAATGATCG
CCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCA
AAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGG
GTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA
GCGGCCGC

OST261

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCTACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA
CTGACGACAGTTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTC
GATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGAGGTGCGGCAGGTTACGTG
TATTAGGTGTCGTCGGGGTTTTCTGTTGACTAGTGGTCCGTGCAAAATGATCG
CCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCA

FIG. 39A

66/89

AAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGG
GTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA
GCGGCCGC

OST262

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA
CTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTC
GATTTCTGGGCAGAGTGGTGCGGTCCGAATTCGGTGTTGGGGGGATGATTTG
GGAGGCTTGGAGTTGTTATGCGTGTGGGACTAGTGGTCCGTGCAAAATGATCG
CCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCA
AAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGG
GTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA
GCGGCCGC

OST263

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCACTGGTGACGGATCCATGAGCGATAAAATTATTCACCTGA
CTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTC
GATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCTTGTGGATTGGGCCGGGTGA
TCAGGGTCTGTTTCGGCGTTTTTGTTTTTTACTAGTGGTCCGTGCAAAATGATCG
CCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCA
AAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGG
GTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA
GCGGCCGC

OST264

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCACTGGTGACGGATCCGTGTCTTCTGATCAAAATCATTTC
GAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACC
TTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGT
GTTGTGCTAGACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGT
CAAGTGTGTACCAAGCATAGGAGAAAAGGCTCTCATGGACTAGAAATATTCCA
GCGTTGTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGAAAGATCACCATC
AAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGAGACACTAAGCGGCCGC

FIG. 39B

67/89

OST265

AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCTACTGGTGACGGATCCTGCGCTAGTCCCACCCGCGGAGGGG
ACGCGGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCTGCATG
CGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTC
TTCTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCT
TTGGTAATGATCATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCT
TCAAAAATGTATCACACCAAAGGACAAGAAGGTTCTGTTTGTCTCCGGTCATC
AGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGTCCAAGATCTGTA
AACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCTCT
CATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCTTGCTA
AGCGGCCCGC

OST266

AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGA
CACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGT
GGTGCGGTCCGAATTCCTATGCGTGTTGTTTTCTTGTAGTAGGTGTAGGTGG
TGGTTGCCTTGGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGA
AATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCAAAACTGAACATCGATC
AAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTG
CTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGG
TCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAAGCGGCCGC

OST267

AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGA
CACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGT
GGTGCGGTCCGAATTCATTTGTGAGGTTGTGAGGTTGTGGAGTCGGTATCCT
TGGTCTTGGGTGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGA
AATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCAAAACTGAACATCGATC
AAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTG
CTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGG
TCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAAGCGGCCGC

FIG. 39C

68/89

OST268

AAGCTTGCCACCATGGGCGATAAAATTATTACCTGACTGACGACAGTTTTGA
CACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGT
GGTGCGGTCCGAATTCGGTTGTACTAGTGCGGTGTGTGGTGCTTGGGCTGAG
GCGGGTAGGTTTTATTGTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCT
GGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCAAACTGAACA
TCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACT
CTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTC
TAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAAGCGGCCGC

OST269 (irrelevant control peptide for OST266-OST268)

AAGCTTGCCACCATGGGCGATAAAATTATTACCTGACTGACGACAGTTTTGA
CACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGT
GGTGCGGTCCGAATTCCTTGTGGATTGGGCCGGGTGATCAGGGTCTGTTTCGG
CGTTTTGTTTTTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGA
AATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCAAACTGAACATCGATC
AAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTG
CTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGG
TCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAAGCGGCCGC

FIG. 39D

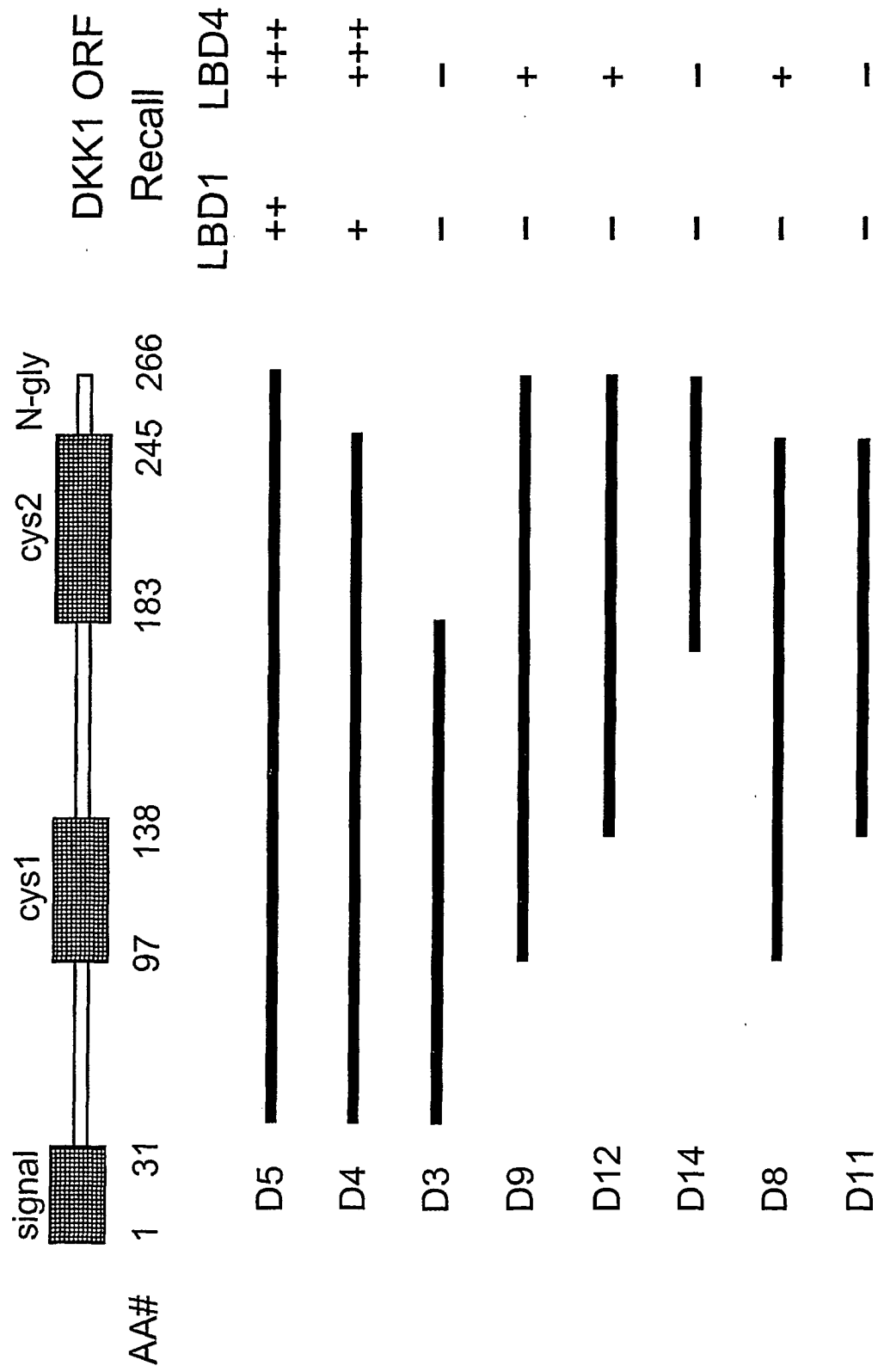


FIG. 40

Effect of Dkk's on Wnt1 signaling
with Coreceptors LRP5, HBM or LRP6
HOB03CE6 Cells

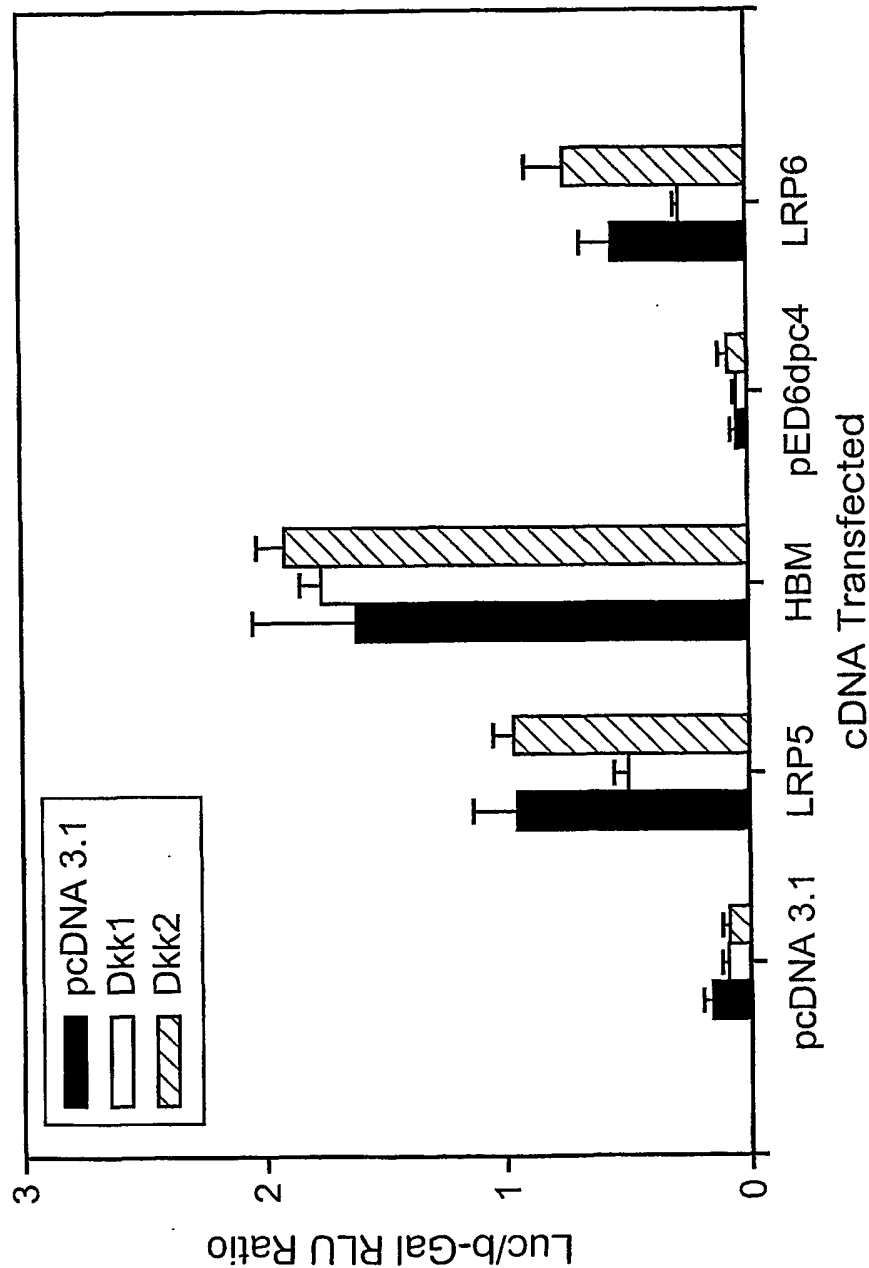


FIG. 41

Effect of Dkk's on Wnt3a Signaling
with Coreceptors LRP5, HBM or LRP6
HOB03CE6 Cells

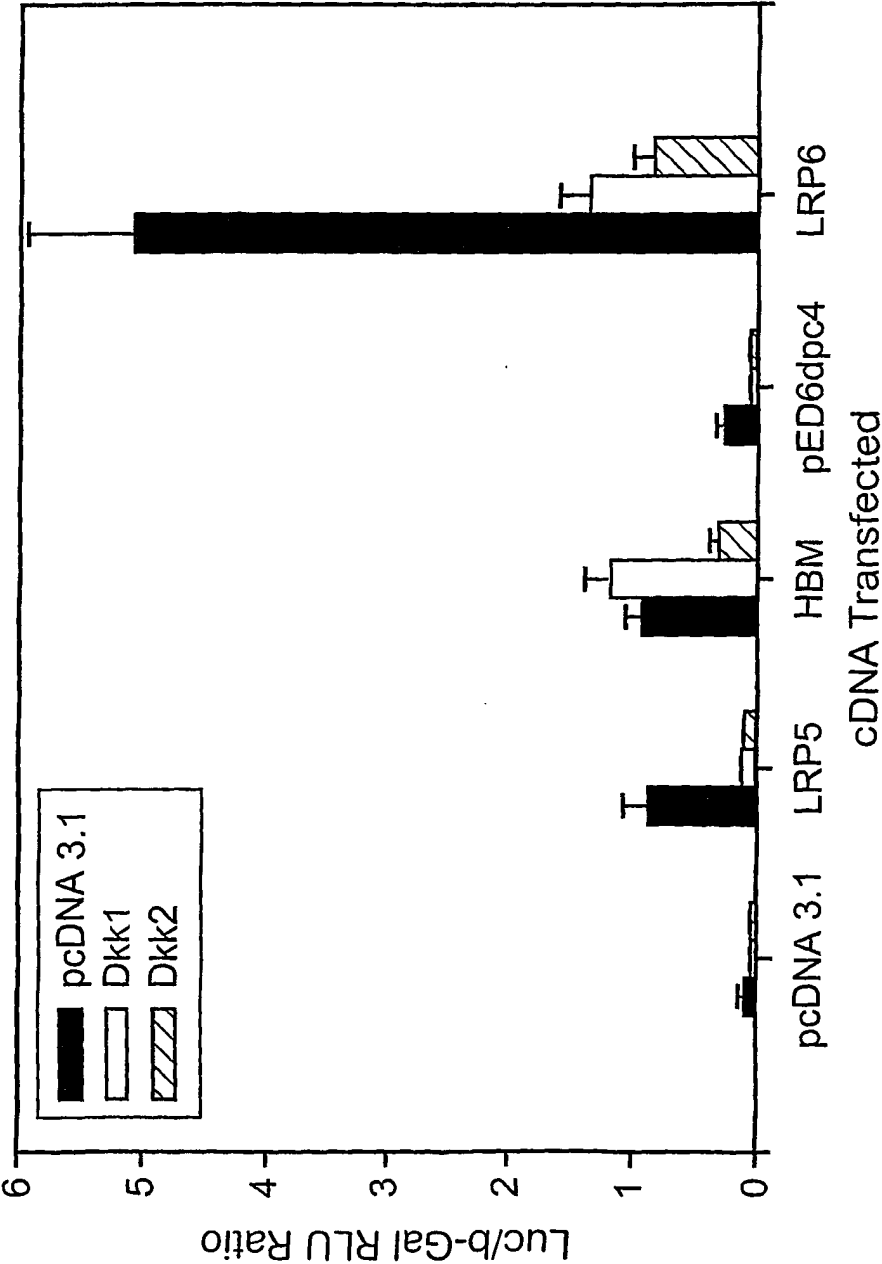


FIG. 42

LRP5-LBD Peptide Aptamer 262 Activates Wnt Signaling in the presence of Wnt3a

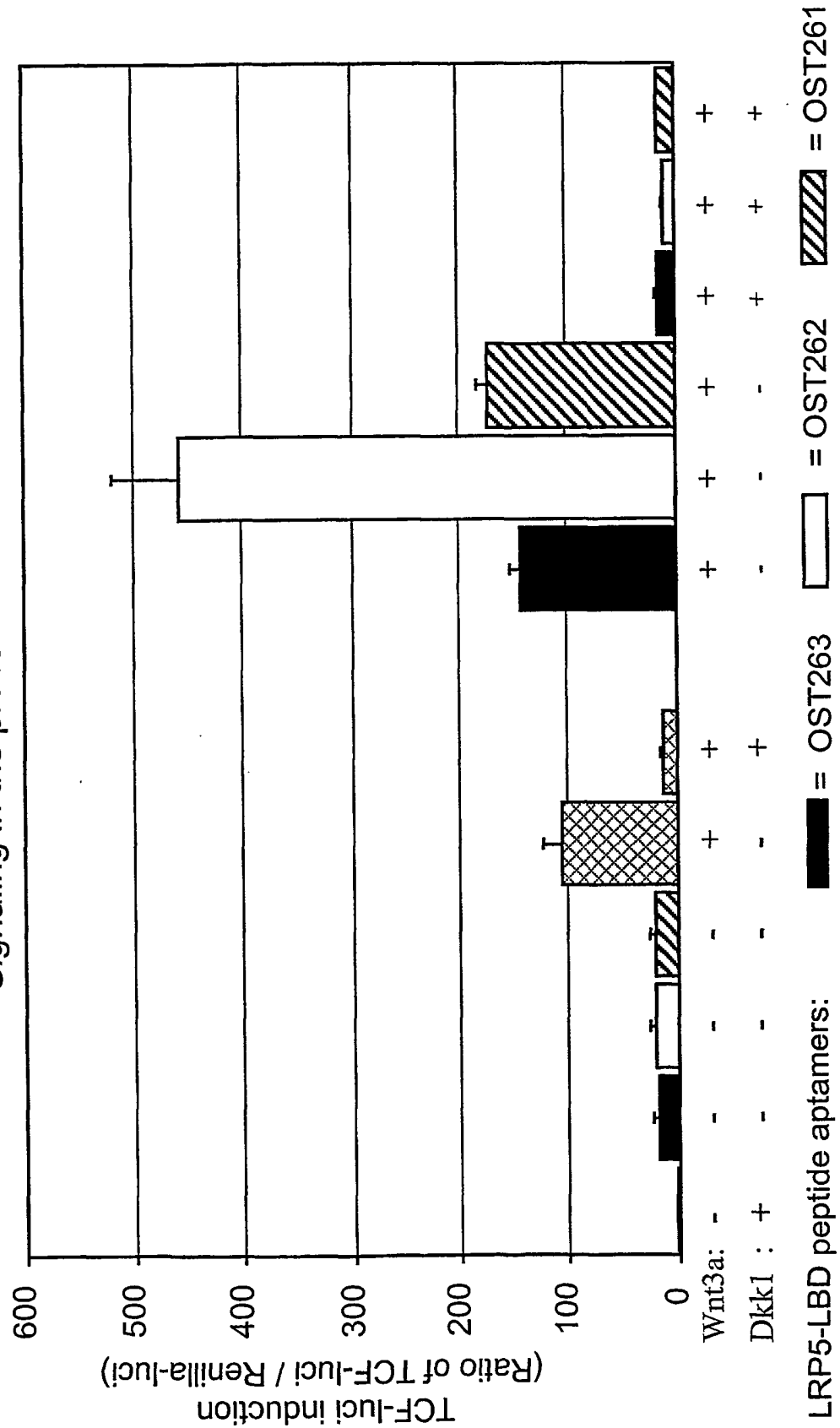


FIG. 43

73/89

Listed below are the amino acid sequences corresponding to the pcDNA3.1 constructs in Appendix 1A

OST258

METDTLLLWVLLLWVPGSTGDGS

OST259

METDTLLLWVLLLWVPGSTGDGSMDSKI IHLTDDSFDTDVLKADGAILVDFWA
EWCGPNSVVLCSRCGRLWRWSCGTSGPCKMIAPILDEIADEYQGKLTVAKLNI
DQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST260

METDTLLLWVLLLWVPGSTGDGSMDSKI IHLTDDSFDTDVLKADGAILVDFWA
EWCGPNSGWRWCGRGALWRRVTS GPCKMIAPILDEIADEYQGKLTVAKLNI
DQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST261

METDTLLLWVLLLWVPGSTGDGSMDSKI IHLTDDSFDTDVLKADGAILVDFWA
EWCGPNSEVRQVTCIRCRRGFLTSGPCKMIAPILDEIADEYQGKLTVAKLNI
DQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST262

METDTLLLWVLLLWVPGSTGDGSMDSKI IHLTDDSFDTDVLKADGAILVDFWA
EWCGPNSGGGMIWEAWSCYACGTSGPCKMIAPILDEIADEYQGKLTVAKLNI
DQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST263

METDTLLLWVLLLWVPGSTGDGSMDSKI IHLTDDSFDTDVLKADGAILVDFWA
EWCGPNSLWIGPGDQGLFRRFVFTSGPCKMIAPILDEIADEYQGKLTVAKLNI
DQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST264

METDTLLLWVLLLWVPGSTGDGSVSSDQNHFRGEIEETITESFGNDHSTLDGY
SRRTTLSSKMYHTKGQEGSVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCT
KHRRKGSHGLEIFQRCYCGEGLScriQKDH HQASNSRLHTCQRH

FIG. 44A

74/89

OST265

METDTLLLLWVLLLLWVPGSTGDGSCASPTRGGDAGVQICLACRKRRKRCMRHAM
CCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMY
HTKGQEGSVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLE
IFQRCYCGEGLSC.

OST266

MGDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPNSYAWLFSCSRCRWLWP
TSGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKN
GEVAATKVGALSKGQLKEFLDANLA

OST267

MGDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPNSICEVRLWSRYPWSWV
TSGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKN
GEVAATKVGALSKGQLKEFLDANLA

OST268

MGDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPNSGCTSAVCGAWAEAGRF
YCTSGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLF
KNGEVAATKVGALSKGQLKEFLDANLA

OST269

MGDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPNSLWIGPGDQGLFRRFVF
TSGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKN
GEVAATKVGALSKGQLKEFLDANLA

FIG. 44B

FIG. 45

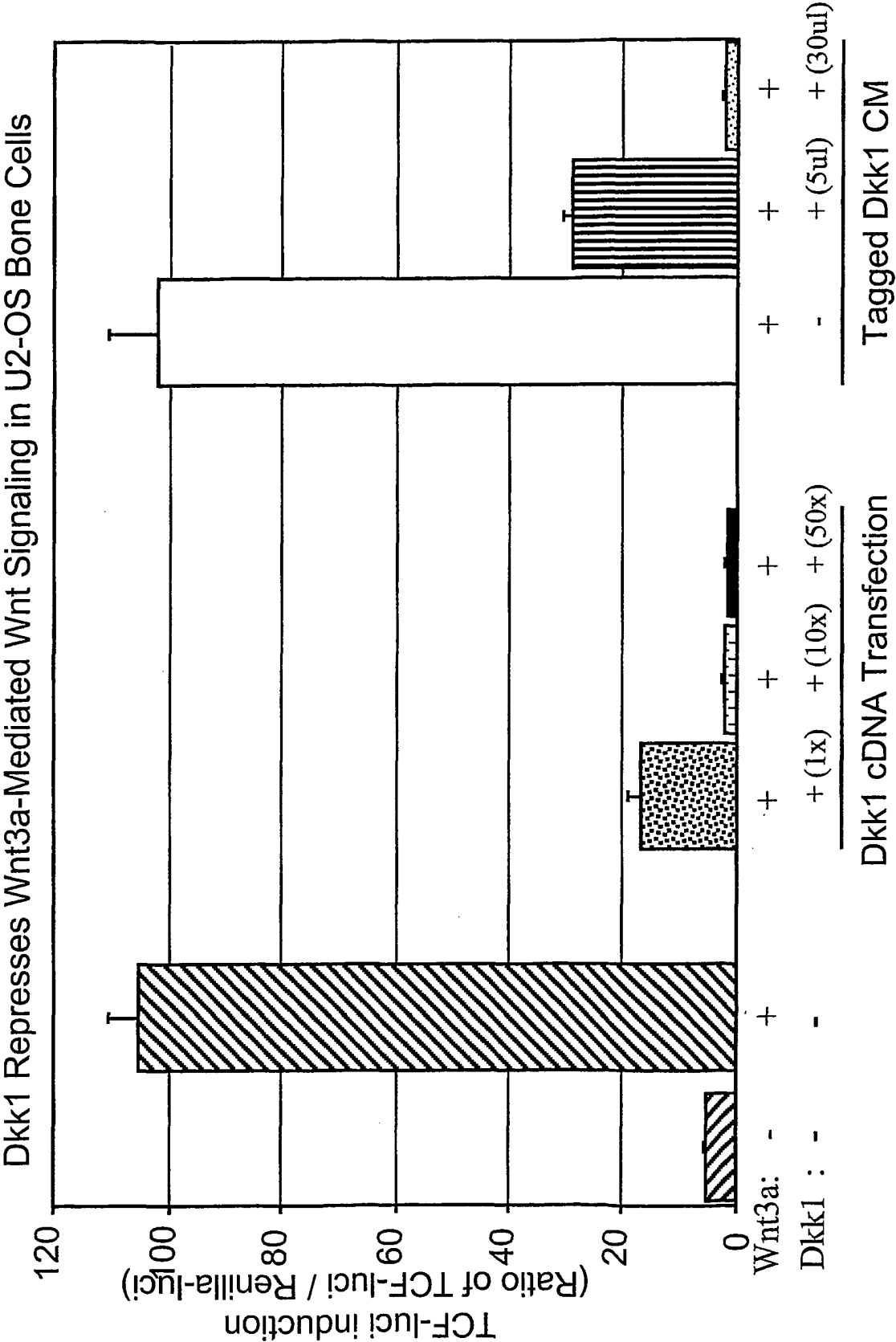
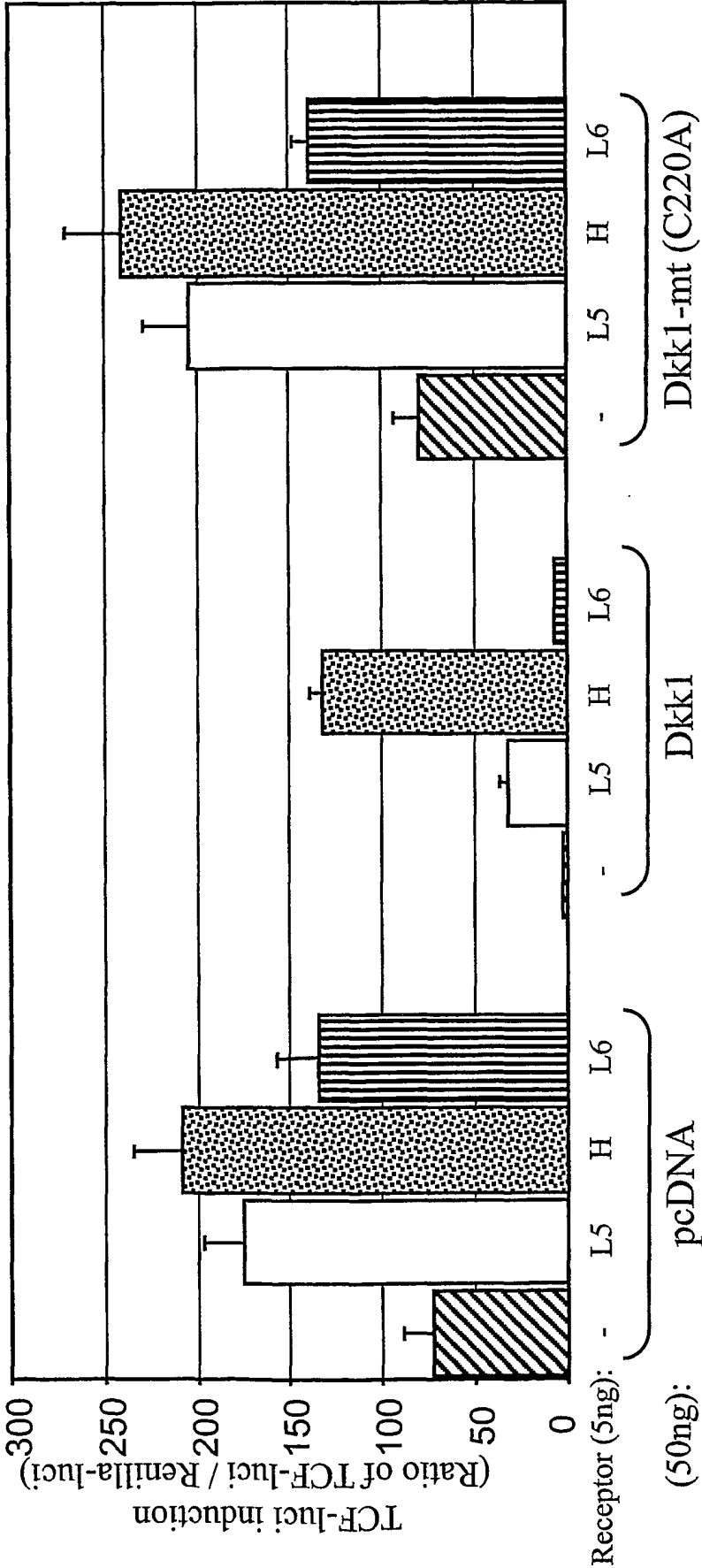


FIG. 46

Wnt1 - HBM generated TCF-luci is not efficiently inhibited by Dkk1 in U2-OS bone cells.



- With Wnt1 the TCF-signal generated by LRP5 is greater than that of LRP6.
- LRP5/6 -Wnt1 induced TCF- is efficiently blocked byDkk1

In U2-OS cells TCF-signal can be modulated by Dkk1, Dkk1-AP,
without Wnt DNA transfection.

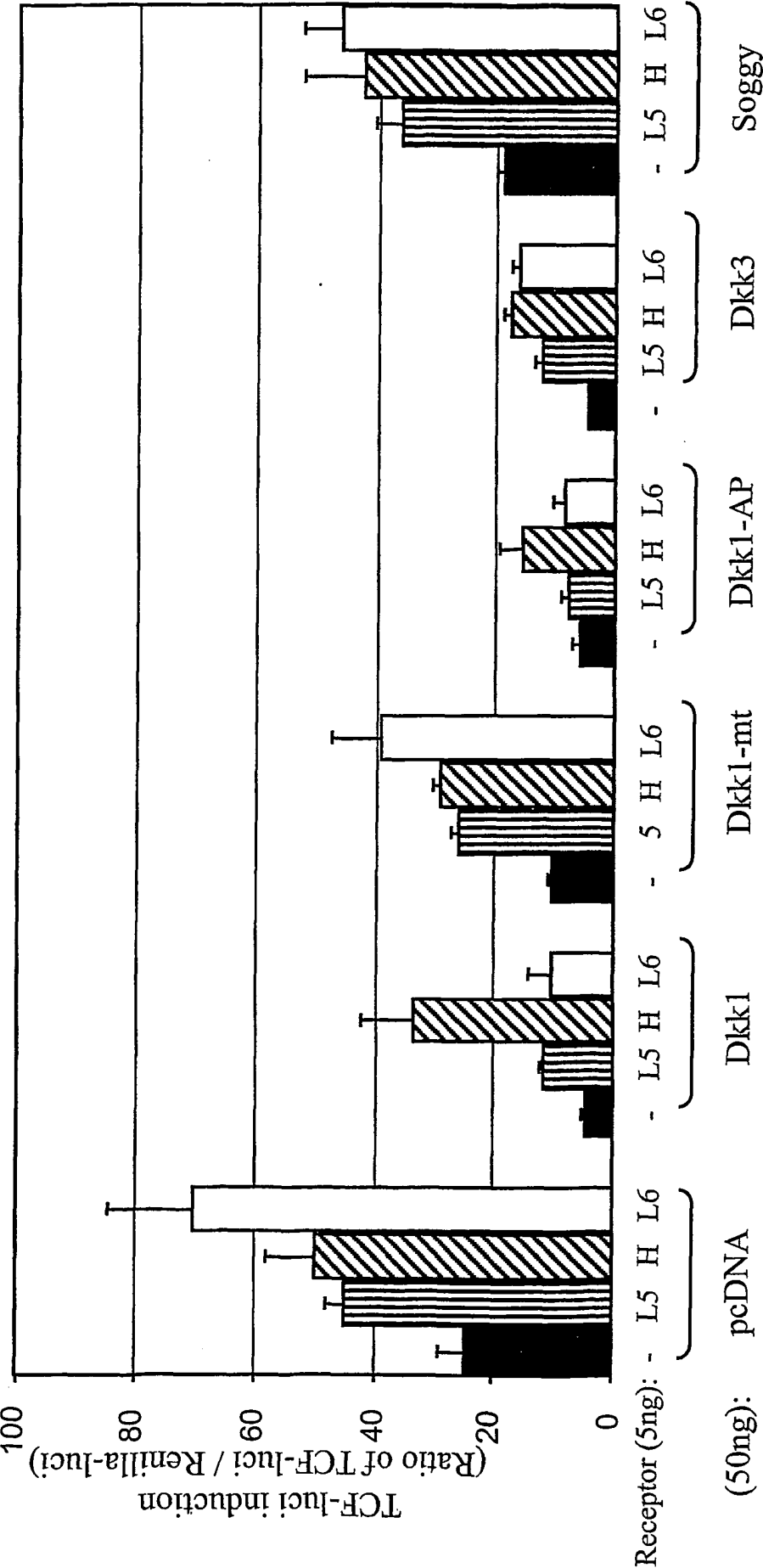


FIG. 47

78/89

FIG. 48

Aptamers 261 and 262 from the LRP5-LBD Activate Wnt Signaling in Xenopus

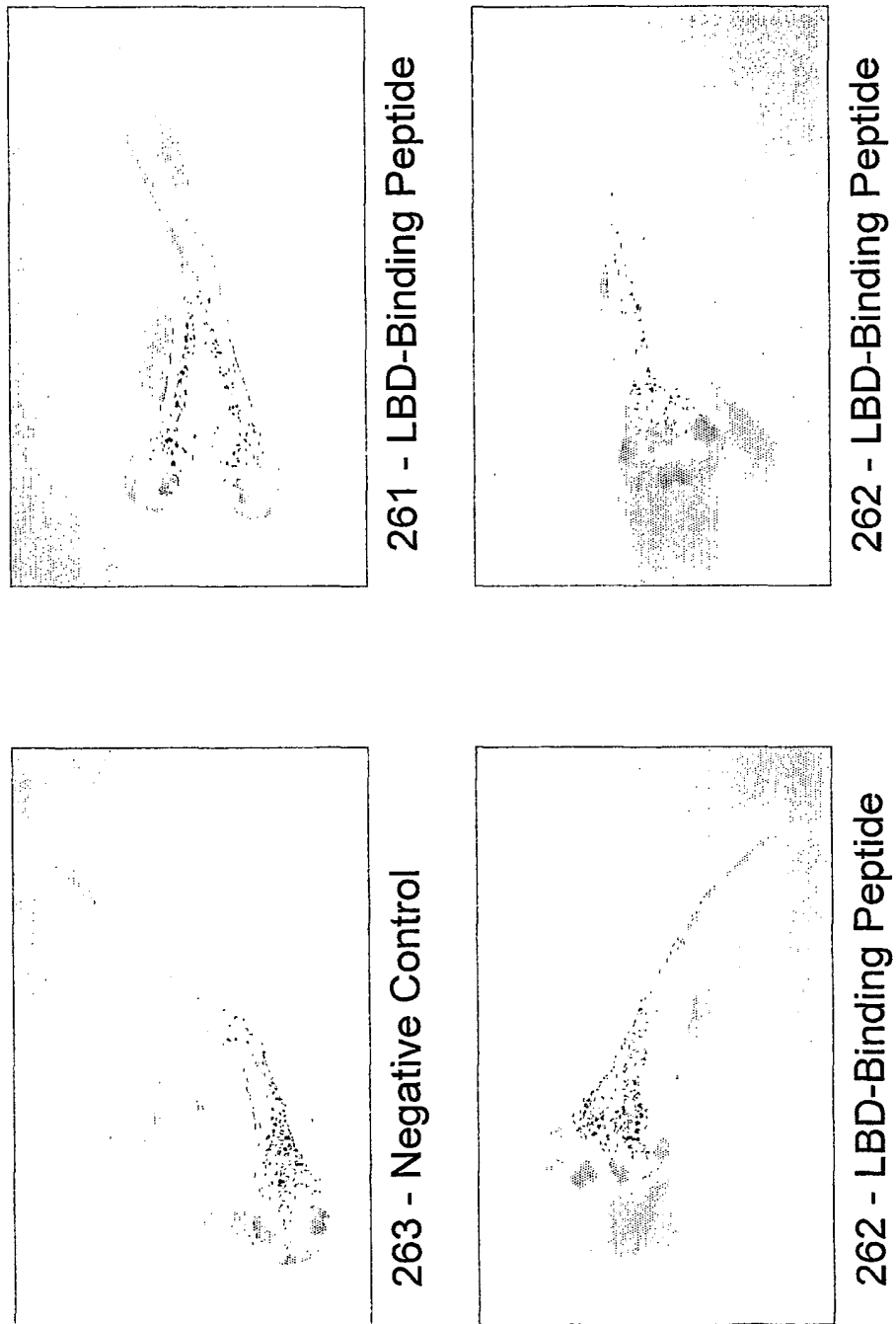
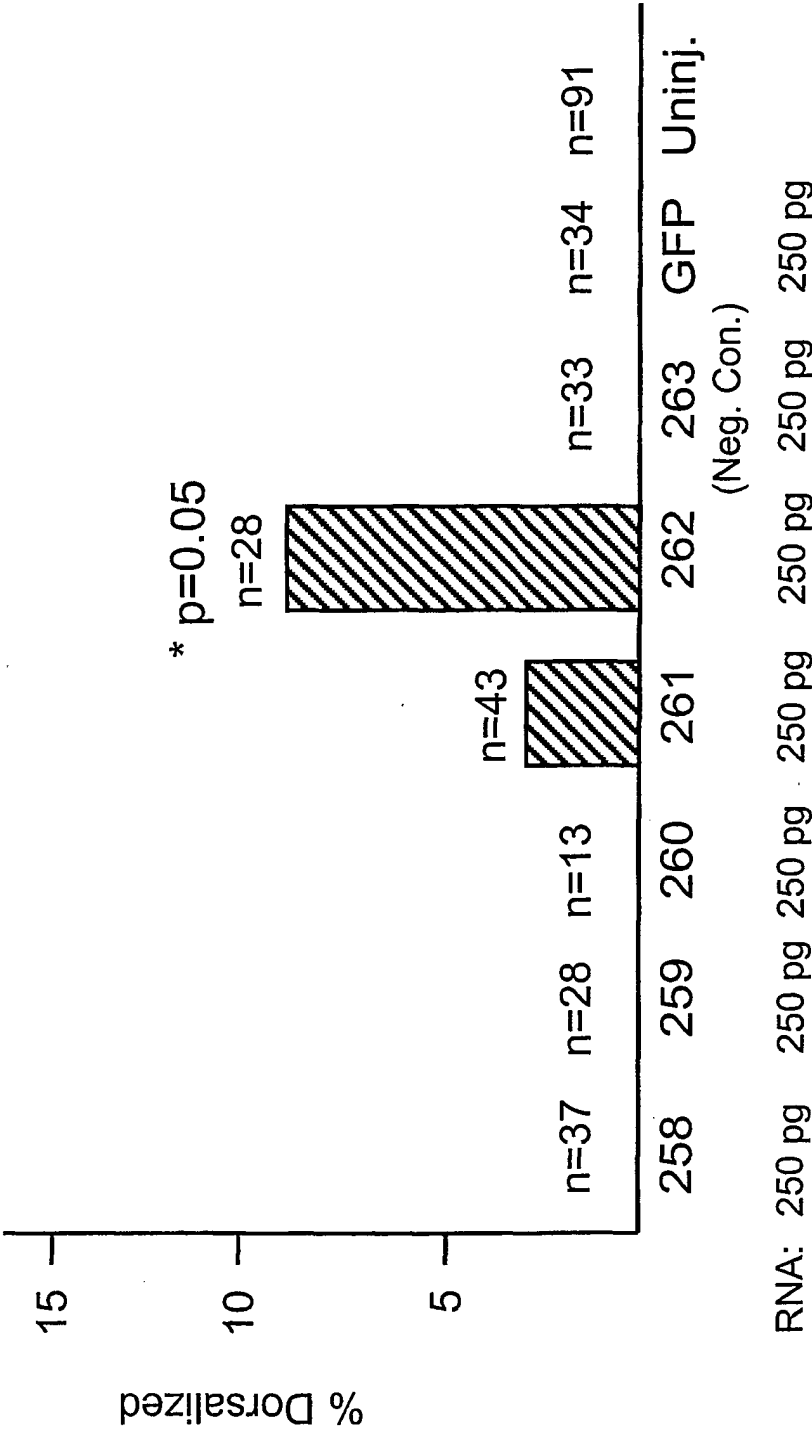


FIG. 49

LRP5 Peptide Aptamers 261 and 262

Induce Wnt Signaling



80/89

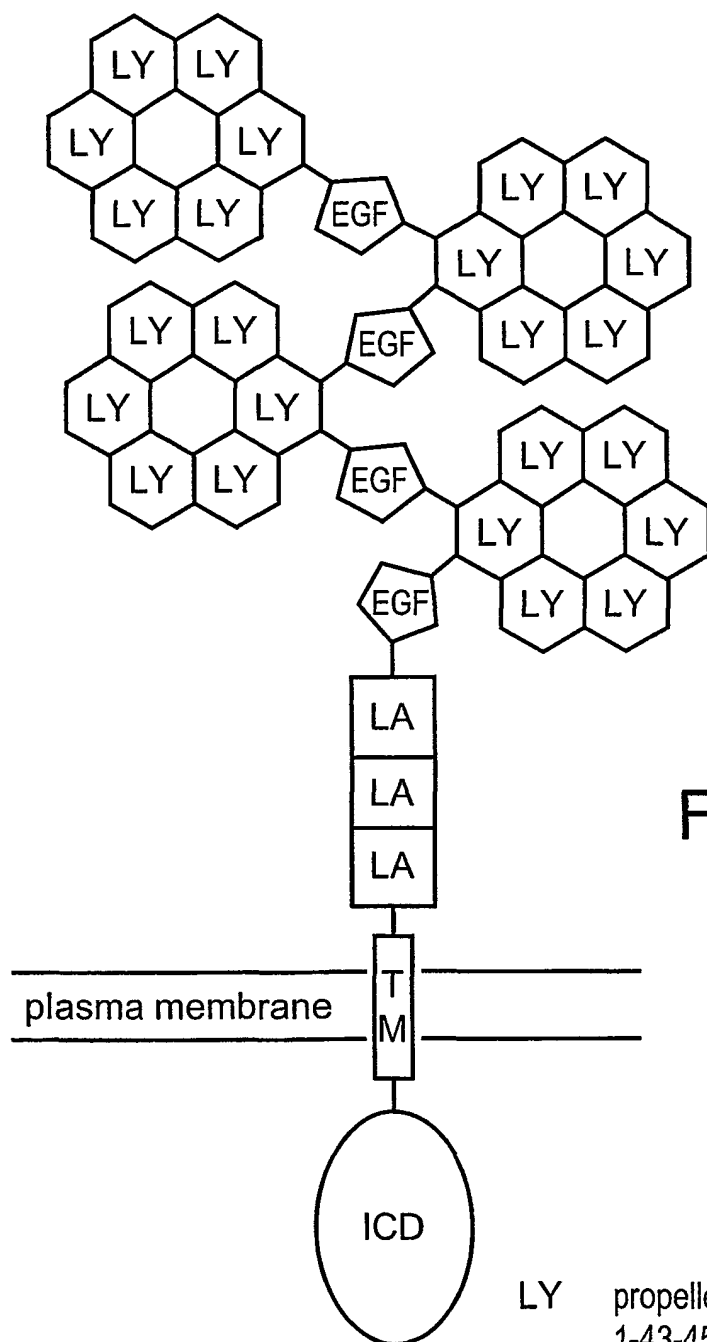


FIG. 50

- LY propeller blade (YWTD "domain"/repeat 1-43-45 aa ae ea./~265/prop.)
- EGF epidermal growth factor-like domain [~42 aa]
- LA LDL receptor class A module [~38 aa]
- TM transmembrane helix [22 aa]
- ICD intracellular domain [107 aa]

Blades 1 @ = **DOUBLE UNDERLINE UPPERCASE** Blades 2 @ = 'lower case – dashed box'

Blades 3 @ = **UPPERCASE BOX** Blades 4 @ = double underline lower case

Blades 5 @ = **UPPERCASE DASHED UNDERLINE** Blades 6 @ = double underline bold lower case

Top loops @ **bold lower case** Bottom loops @ **BOLD UPPER CASE**

AF064984m_ywtd6a	-ASPLLLFANRRDVRLVDAGGVKL-----ESTIVASGLEDAAAVDFQFSKGAVYWTDVSEE
AF077847m_ywtd6a	-ASPLLLFANRRDVRLVDAGGVKL-----ESTIVASGLEDAAAVDFQFSKGAVYWTDVSEE
AF077820h_ywtd6a	-ASPLLLFANRRDVRLVDAGGVKL-----ESTIVSGLLEDAAAVDFQFSKGAVYWTDVSEE
AF074264h_ywtd6a	--APLLLYANRRDLRLVDATNGKE-----NATIVVGGLEDAAAVDFVFSHGGLIYWSDVSEE
AF074265m_ywtd6a	--APLLLYANRRDLRLVDATNGKE-----NATIVVGGLEDAAAVDFVFGHGLIYWSDVSEE
11px	DEEPFLIFANryYLRKLNLDGS-----NYTLKQglnnavalidfDYREQmi_ywtdvttg
af064984m_ywtd6c	-PEAFLVFTSRATIHRIISLETNNN-----DVAIPLTGVKEASALDFDVSNNNHIYWTDVSLK
af077820h_ywtd6c	-PEAFLVFTSRAAIHRIISLETNNN-----DVAIPLTGVKEASALDFDVSNNNHIYWTDVSLK
af074264h_ywtd6c	-PEAFLLFSRRADIRRIISLETNNN-----NVAIPLTGVKEASALDFDVTDNRIYWTDISLK
af074265m_ywtd6c	-PEAFLLFSRRADIRRIISLETNNN-----NVAIPLTGVKEASALDFDVTDNRIYWTDISLK
af064984m_ywtd6b	GAEVLLARRTDLRRIISLDTPDF-----TDIVLQVGDIRHAIADYDPLEGYVYWTDDEV
af077820h_ywtd6b	---EVLLLARRTDLRRIISLDTPDF-----TDIVLQVDDIRHAIADYDPLEGYVYWTDDEV
af074264h_ywtd6b	GATELLLLARRTDLRRIISLDTPDF-----TDIVLQLEDIRHAIADYDPPVEGYIYWTDDEV
af074265m_ywtd6b	GATELLLLARRTDLRRIISLDTPDF-----TDIVLQLEDIRHAIADYDPPVEGYIYWTDDEV
1ndx	PPGTHLLFAQtgKIERLPLEGNTMRKTEAKAFHvpakvigi lafbcVDKmmvwt d i tep
af064984m_ywtd6d	-PSTFLLFSQKFAISRMIPDDQLS-----PDLVLPLHGLRNVKAINYDPLDKFIYWVDGRQN
af077820h_ywtd6d	-PTTFLLFSQKSAISRMIPDDQHS-----PDLILPLHGLRNVKAIDYDPLDKFIYWVDGRQN

FIG. 51A

af074264h_ywtd6d
af074265m_ywtd6d
DSC_SEC
PROB_H
PROB_E
PROB_C
AF064984m_ywtd6a
AF077847m_ywtd6a
AF077820h_ywtd6a
AF074264h_ywtd6a
AF074265m_ywtd6a
11px
af064984m_ywtd6c
af077820h_ywtd6c
af074264h_ywtd6c
af074265m_ywtd6c
af064984m_ywtd6b
af077820h_ywtd6b
af074264h_ywtd6b
af074265m_ywtd6b
1idx
af064984m_ywtd6d
af077820h_ywtd6d
af074264h_ywtd6d

-PTTFLFSQKSAINRMVIDEQQS---PDIILPIHSLRNVRAIDYDPLDKQLYWIDSRQN
-PSTFLLFSQKSAINRMVIDEQQS---PDIILPIHSLRNVRAIDYDPLDKQLYWIDSRQN
-CCCEHHHHCCCCCEEECCCCC-----CEEECCCCCEEECCCCCEEECCCCCHH
-112435985444432121110-----0000024344431000111111000011257
-00256301010111665321111-----299999611011569975422135888632211
-99611211556555124578889-----81111256655231134577864222467642
AIKQTYLNQTG---AAQNIIVISGLVSPDGLACDWVGKLYWTDSETNRIEVANLNGTS
AIKQTYLNQTG---AAQNIIVISGLVSPDGLACDWVGKLYWTDSETNRIEVANLNGTS
AIKQTYLNQTG---AAQNVVVISGLVSPDGLACDWVGKLYWTDSETNRIEVANLNGTS
AIKRTEFNKT-----ESVQNVVVSGLLSPDGLACDWLGEKLYWTDSETNRIEVSNLDGSL
AIKRTEFNKS-----ESVQNVVVSGLLSPDGLACDWLGEKLYWTDSETNRIEVSNLDGSL
gsmiirmhING-----snvqvlhrtg1snPDGLAVDPWVGGLYWCdkgrdTIEVSKLNGAY
TISRAFMNGS-----SVEHVIEFFGLDYPEGMAVDWMGKNLYWADTGTNRIEVARLDGQF
TISRAFMNGS-----SVEHVVEFFGLDYPEGMAVDWMGKNLYWADTGTNRIEVARLDGQF
TISRAFMNGS-----ALEHVVEFFGLDYPEGMAVDWLGNLYWADTGTNRIEVSCLDGGH
TISRAFMNGS-----ALEHVVEFFGLDYPEGMAVDWLGNLYWADTGTNRIEVSCLDGGH
AIRRAYLDGS-----GAQTLVNTEINDPDGIAVDWVARNLYWTDGTDRIEVTRLNGTS
AIRRAYLDGS-----GAQTLVNTEINDPDGIAVDWVARNLYWTDGTDRIEVTRLNGTS
AIRRSFIDGS-----GSQFVVTAQIAHPDGIADVWVARNLYWTDGTDRIEVTRLNGTM
AIRRSFIDGS-----GSQFVVTAQIAHPDGIADVWVARNLYWTDGTDRIEVTRLNGTM
sigrasLHGG-----EpttliinqdlgsPEGIAVDHLGRNIFWTDsnldrIEVAKLDGTQ
-IKRAKDDGTQPS-MLTSPSQSLSPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLDGDA
-IKRAKDDGTQPF-VLTSLSQGNPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEA
MIRKAQEDGSQGFTVVVSSVPSONLEIQPYDLSIDIYSRYIYWTCEATNVINVTRLDGRS

82/89

FIG. 51B

83/89

af074265m_ywtd6d
DSC_SEC
PROB_H
PROB_E
PROB_C

SIRKAHEDGGQGFNVVANSVANQNLEIQPYDLSIDIYSRYIYWTCEATNVIDVTRLDGSR
HHHHEEECCCC---CC
99953211111---113100001111111001332220001111110000221112
00135652111---123799996421111257965323367751100169999630001
11022247888---874211113578888743132465543348899831111259997

RKVLFWQDLDPRAIALDPAHGYMYWTDWG-EAPRIERAGMDGSTRKIIIVDSDIYWPNGL
RKVLFWQDLDPRAIALDPAHGYMYWTDWG-EAPRIERAGMDGSTRKIIIVDSDIYWPNGL
RKVLFWQDLDPRAIALDPAHGYMYWTDWG-ETPRIERAGMDGSTRKIIIVDSDIYWPNGL
RKVLFWQELDPRAIALDPSSGFMYWTDWG-EVPKIERAGMDGSSRFIIINSEIYWPNGL
RKVLFWQELDPRAIALDPSSGFMYWTDWG-EVPKIERAGMDGSSRFVINTIYWPNGL
RTVLVNsglrePRALVVDVQNGLYWTDwg-dhsLIGKIGMDGTNRSVIVdtkitwPNGL
RQVLVWRDLNPRSLALDPTKGYIYWTEWG-GKPRIVRAFMDDGTNCMTLVD-KVGRANDL
RQVLVWRDLNPRSLALDPTKGYIYWTEWG-GKPRIVRAFMDDGTNCMTLVD-KVGRANDL
RQVLVWKDLSPRALALDPAEGFMYWTEWG-GKPKIDRAAMDGSERTTLVP-NVGRANGL
RQVLVWKDLSPRALALDPAEGFMYWTEWG-GKPKIDRAAMDGSERTTLVP-NVGRANGL
RKILVSEDLDEPRAIIVLHPVMGLMYWTDWG-ENPKIECANLDGDRHVLVNTSLGWPNGL
RKILVSEDLDEPRAIIVLHPVMGLMYWTDWG-ENPKIECANLDGQERRVLVNASLGWPNGL
RKILISEDLEPRAIIVLDPVMGYMYWTDWG-EIPKIERAALDGSDRVVLVNTSLGWPNGL
RKILISEDLEPRAIIVLDPVMGYMYWTDWG-EIPKIERAALDGSDRVVLVNTSLGWPNGL
RRVLfetdlvnPRGIVTDSVRGNLYWTDwnrdnpKIETSYMDGTNRRIIVgddlglPNGL
MGVVLRGDRDKPRAIAVNAERGMYFTNMQDHAAKIERASLDGTEREVLTFTGLIRPVAL
MGVVLRGDRDKPRAIVVNAERGMYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVAL
VGVVLKGEQDRPRAIVVNPEKGYMYFTNLQERSPKIERAALDGTREVLFFSGLSKPIAL

AF064984m_ywtd6a
AF077847m_ywtd6a
AF077820h_ywtd6a
AF074264h_ywtd6a
AF074265m_ywtd6a

1lp
af064984m_ywtd6c
af077820h_ywtd6c
af074264h_ywtd6c
af074265m_ywtd6c
af064984m_ywtd6b
af077820h_ywtd6b
af074264h_ywtd6b
af074265m_ywtd6b

1nd
af064984m_ywtd6d
af077820h_ywtd6d
af074264h_ywtd6d

FIG. 51C

af074265m_ywtd6d
DSC_SEC
PROB_H
PROB_E
PROB_C
AF064984m_ywtd6a
AF077847m_ywtd6a
AF077820h_ywtd6a
AF074264h_ywtd6a
AF074265m_ywtd6a
1lpx
af064984m_ywtd6c
af077820h_ywtd6c
af074264h_ywtd6c
af074265m_ywtd6c
af064984m_ywtd6b
af077820h_ywtd6b
af074264h_ywtd6b
af074265m_ywtd6b
1ndx
af064984m_ywtd6d
af077820h_ywtd6d
af074264h_ywtd6d
af074265m_ywtd6d
DSC_SEC

VGVLKGEQDRPRAIVVNPEKGYMYFTNLQERSPKIERAALDGTEREVLFFSGLSKPIAL
CEEEEECCCCCCCCCCCCCCCC-CCCHHHHCCCCCCCCCCCCCCCCCCCC
331102433213441111221110001111-1125475642111220000000011112
359996111001117651100159996200-00123232110001499998222111123
420012566896552348789841113799-9973312257999741111288898875
TIDLEEQLYWADAKLSFIHRANLDGSFRQKVVEGSLTHPFALTLSGDTLYWTDWQTRSI
TIDLEEQLYWADAKLSFIHGANLDGSFRQKVVEGSLTHPFALTLSGDTLYWTDWQTRSI
TIDLEEQLYWADAKLSFIHRANLDGSFRQKVVEGSLTHPFALTLSGDTLYWTDWQTRSI
TIDLEEQLYWADAKLSFIHKSNDGTNRQAVVKGSLPHPFALTLFEDILYWTDWSTHSI
TIDYQERKLYWADAKLNFHKSNDGTNRQAVVKGSLPHPFALTLFEDTLYWTDWNTHSI
TIDYINSRIYWADaredYIEFASLDGNSRHTVLSqdi phi faltlFEDyiywtdwetksi
TIDYADQRLYWTDLDTNMISSNMLGQER-MVIADDLPPFGLTQYSDYIYWTDWNLHSI
TIDYADQRLYWTDLDTNMISSNMLGQER-VVIADDLPHPFGLTQYSDYIYWTDWNLHSI
TIDYAKRRRLYWTDLDTNLISSNMLGLNR-EVIADDLPHPFGLTQYQDYIYWTDWSRRSI
TIDYAKRRRLYWTDLDTNLISSDMLGLNR-EVIADDLPHPFGLTQYQDYIYWTDWSRRSI
ALDLQEGKLYWGDAKTDKIEVINIDGTRKRTLLEDKLPHIFGFTLLGDFIYWTDWQRRSI
ALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHIFGFTLLGDFIYWTDWQRRSI
ALDYDEGKIYWGDAKTDKIEVMNTDGTGRRVLVEDKIPHIFGFTLLGDYVYWTDWQRRSI
ALDYDEGTIYWGDAKTDKIEVMNTDGTGRRVLVEDKIPHIFGFTLLGDYVYWTDWQRRSI
HFD AFSSQLCWVDagtnRAECLNPSQPSRRKALeg-lqypfavtSYGknlyftdwkmsv
VVDNALGKLFWVDADLKRIESCDLSGANRLTLEDANIVQPVGLTVLGRHLYWIDRQQQMI
VVDNTLGKLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLWIDRQQQMI
ALDSRLGKLFWADSDLLRRIESSDLSGANRIVLEDSNIIQPVGLTVFFENWLYWIDKQQQMI
ALDSKLGKLFWADSDLLRRIESSDLSGANRIVLEDSNIIQPVGLTVFFENWLYWIDKQQQMI
CCCHHH

FIG. 51D

85/89

PROB_H 11344433100111124433211211224110121111111002112211112224577
 PROB_E 443111137995221215556521000114896321111245663211366642110111
 PROB_C 55455564211477865122247799875511357888875445787533356776422
 AF064984m_ywtd6a HACNKWTGEQRKEILSALYSPMDIQVLSQERQPPFHT
 AF077847m_ywtd6a HACNKWTGEQRKEILSALYSPMDIQVLSQERQPPFHT
 AF077820h_ywtd6a HACNKRTGGRKEILSALYSPMDIQVLSQERQPPFHT
 AF074264h_ywtd6a LACNKYTGEGLREIHSDFSPMDIHAFSQQRQPN--
 AF074265m_ywtd6a LACNKYTGEGLREIHSNIFSPMDIHAFSQQRQPN--
 11px nrhktTGanktllistlhrPMDIHlYHPYR-----
 af064984m_ywtd6c ERADKTSGRNRRTLIQGHLDVMDILVFSRRQDGL--
 af077820h_ywtd6c ERADKTSGRNRRTLIQGHLDVMDILVFSRRQDGL--
 af074264h_ywtd6c ERANKTSGQNRRTIIQGHLDVMDILVFSRRQSG--
 af074265m_ywtd6c ERANKTSGQNRRTIIQGHLDVMDILVFSRRQAG--
 af064984m_ywtd6b ERVHKVK-ASRDVIDQLPDLMLGLKAVNAKVVG--
 af077820h_ywtd6b ERVHKVK-ASRDVIDQLPDLMLGLKAVNAKVVG--
 af074264h_ywtd6b ERVHKRS-AEREVIDQLPDLMLGLKATNVHRVIG--
 af074265m_ywtd6b ERVHKRS-AEREVIDQLPDLMLGLKATSVHRVIG--
 1ndx valdAIsketdafqphkqtrLYGITALSQCP-----
 af064984m_ywtd6d ERVEKTTGDKRTRVQGRVTHLTGIHAVEEVSLEEFSA
 af077820h_ywtd6d ERVEKTTGDKRTRIQGRVAHLTGIHAVEEVSLEEFSA
 af074264h_ywtd6d EKIDMTGREGRTKVQARIAQLSDIHAVKELNLQEYR-
 af074265m_ywtd6d EKIDMTGREGRTKVQARIAQLSDIHAVKELNLQEYR-
 DSC_SEC HHHCCCCCCCCCEEEEECCCCCEEEEECCCCCCCCC
 PROB_H 76644221123211223222211112332211101
 PROB_E 121112100136886432212256666332122111
 PROB_C 2235567997421124466767433324467778898

FIG. 51E

Possible Effect of Functional Mutations on Side Chain Interactions

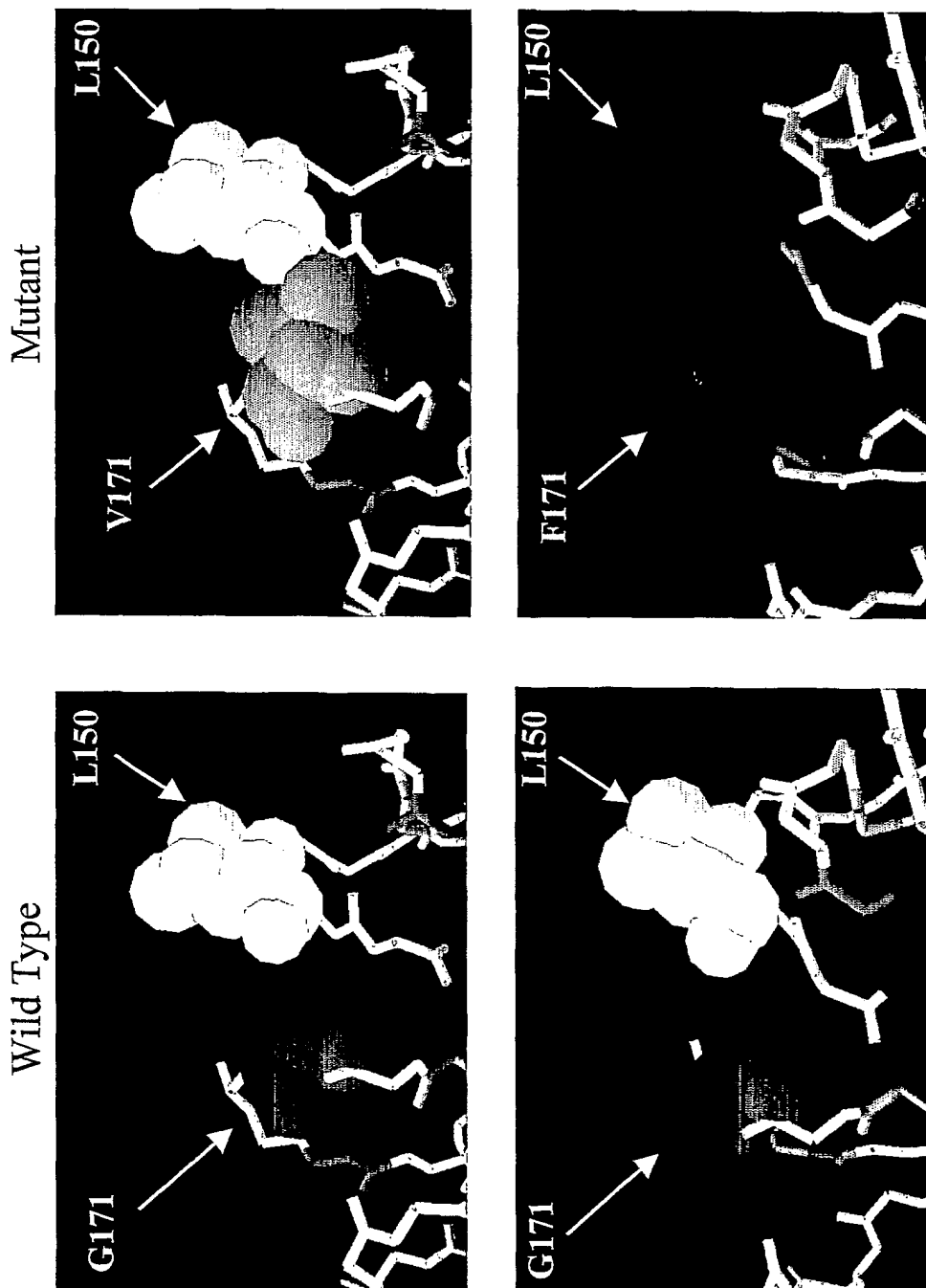


FIG. 52

Homology Model of LRP5's 2nd Propeller Domain
Propeller 2 Contains Two OPPG mutations

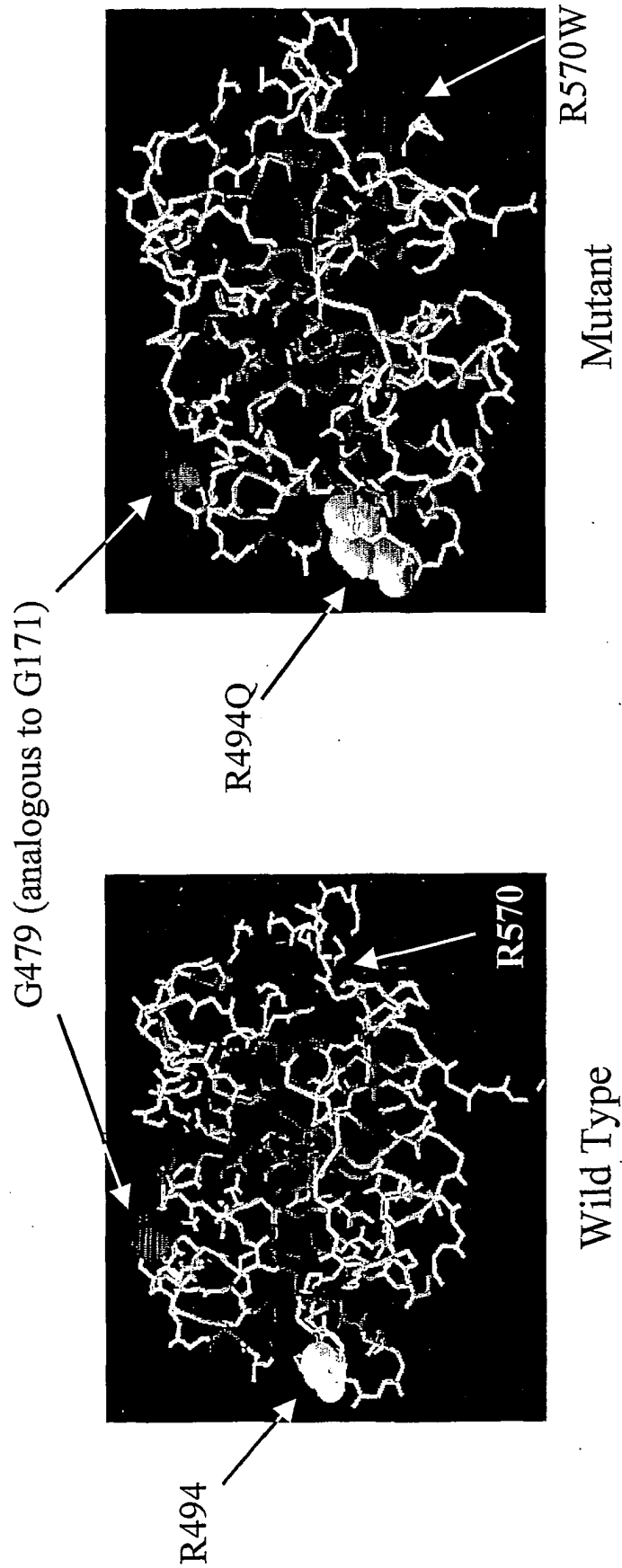


FIG. 53A

88/89

Homology Model of LRP5's 2nd Propeller Domain
Propeller 2 Contains Two OPPG mutations

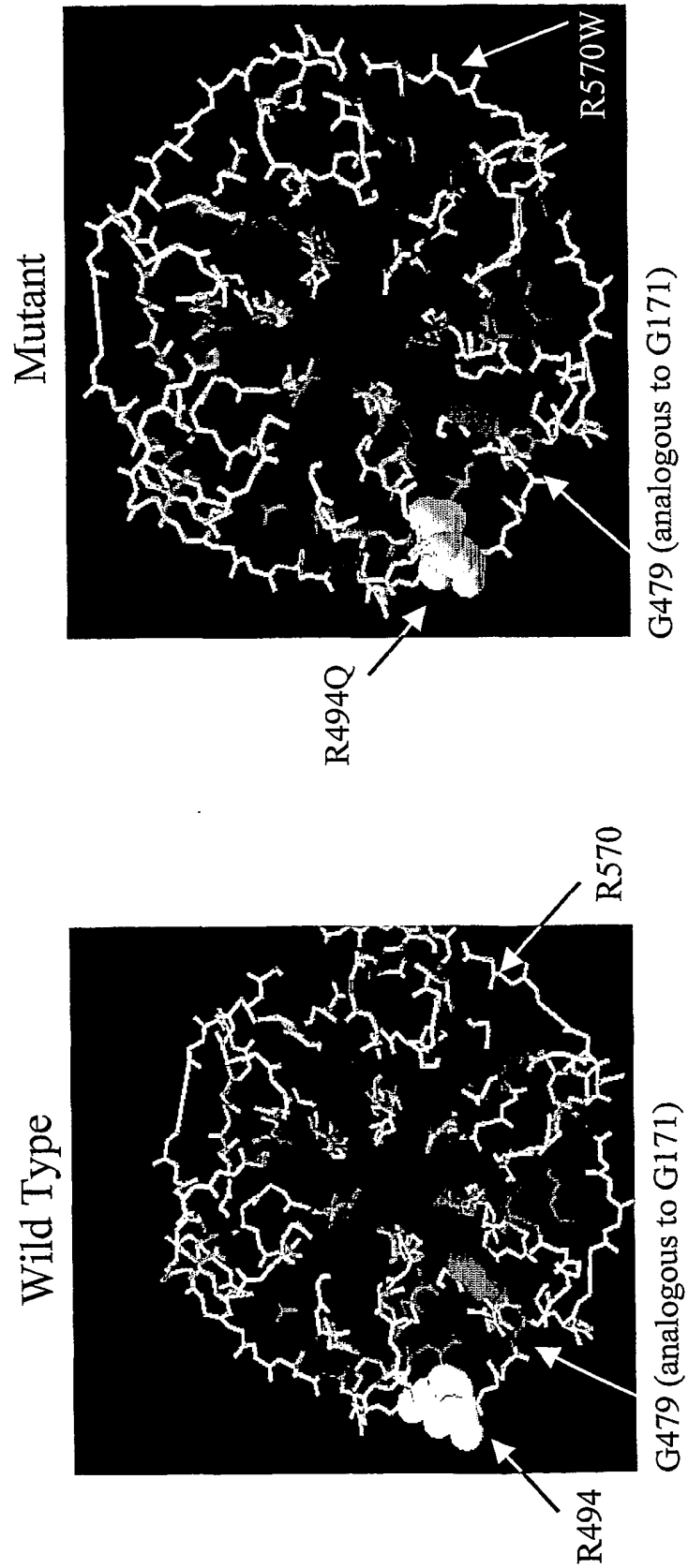


FIG. 53B

89/89

Homology Model of LRP5's 3rd Propeller Domain
Propeller 3 Contains One OPPG mutations

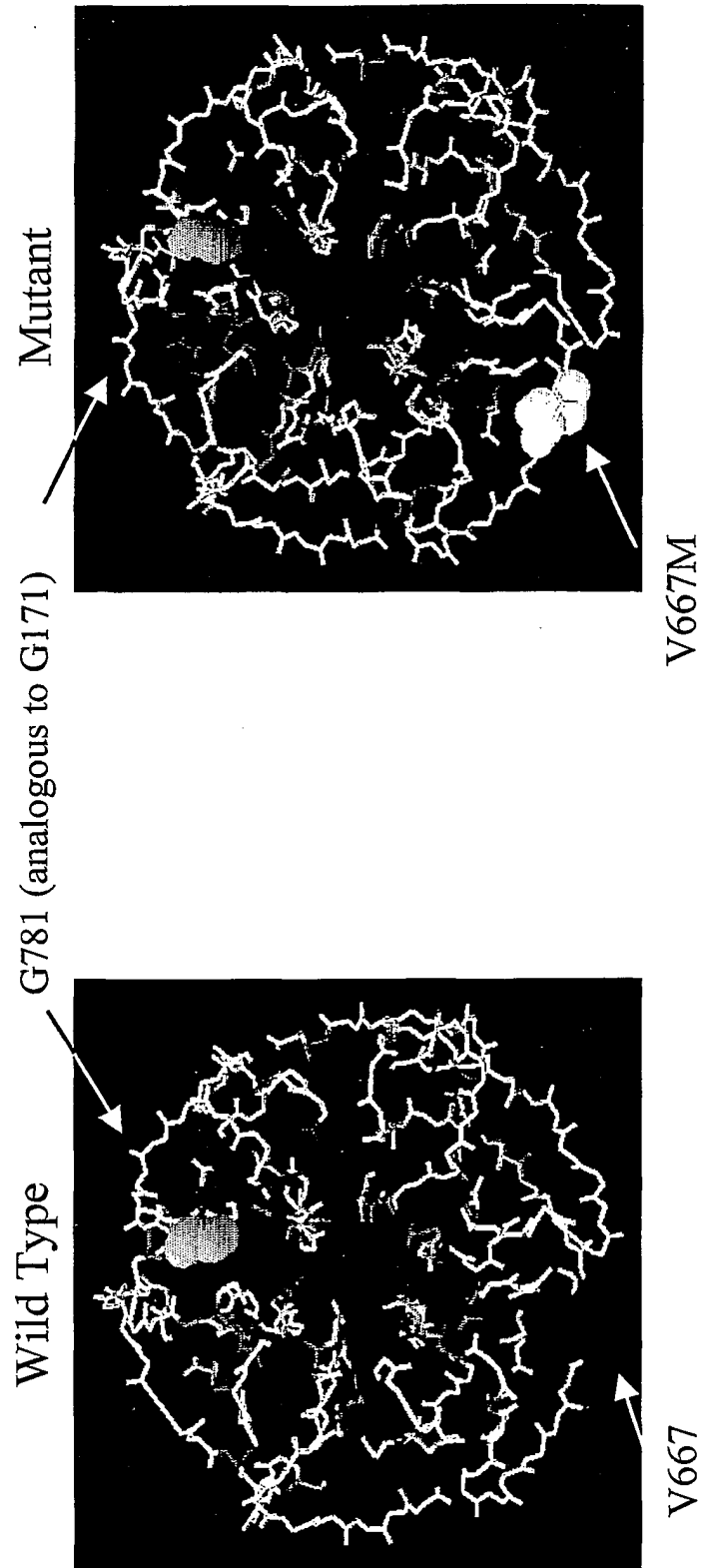


FIG. 53C

SEQUENCE LISTING

<110> Allen, Kristina M.
 Yaworsky, Paul
 Morales, Arturo J.
 Graham, James R.
 Anisowicz, Anthony
 Liu, Wei

<120> HBM Variants that Modulate Bone Mass and Lipid Levels

<130> 032796-135

<150> US 60/290,071

<151> 2001-05-11

<150> US 60/291,311

<151> 2001-05-17

<150> US 60/353,058

<151> 2002-02-01

<150> US 60/361,293

<151> 2002-03-04

<160> 765

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 5120

<212> DNA

<213> Homo sapiens

<400> 1

```

actaaagcgc cgccgcgcgcg ccatggagcc cgagtgcgcg cggcgcgggc ccgtccggcc      60
gccggacaac  atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg      109
              Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu
              1          5          10
ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc      157
Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala
              15          20          25
gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg      205
Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu
              30          35          40          45
gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc      253
Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly
              50          55          60
ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg      301
Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val
              65          70          75
tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac      349
Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn

```

	80					85					90					
cag	acg	ggg	gcc	gcc	gtg	cag	aac	gtg	gtc	atc	tcc	ggc	ctg	gtc	tct	397
Gln	Thr	Gly	Ala	Ala	Val	Gln	Asn	Val	Val	Ile	Ser	Gly	Leu	Val	Ser	
	95						100				105					
ccc	gac	ggc	ctc	gcc	tgc	gac	tgg	gtg	ggc	aag	aag	ctg	tac	tgg	acg	445
Pro	Asp	Gly	Leu	Ala	Cys	Asp	Trp	Val	Gly	Lys	Lys	Leu	Tyr	Trp	Thr	
110							115				120				125	
gac	tca	gag	acc	aac	cgc	atc	gag	gtg	gcc	aac	ctc	aat	ggc	aca	tcc	493
Asp	Ser	Glu	Thr	Asn	Arg	Ile	Glu	Val	Ala	Asn	Leu	Asn	Gly	Thr	Ser	
				130					135					140		
cgg	aag	gtg	ctc	ttc	tgg	cag	gac	ctt	gac	cag	ccg	agg	gcc	atc	gcc	541
Arg	Lys	Val	Leu	Phe	Trp	Gln	Asp	Leu	Asp	Gln	Pro	Arg	Ala	Ile	Ala	
				145					150					155		
ttg	gac	ccc	gct	cac	ggg	tac	atg	tac	tgg	aca	gac	tgg	ggt	gag	acg	589
Leu	Asp	Pro	Ala	His	Gly	Tyr	Met	Tyr	Trp	Thr	Asp	Trp	Gly	Glu	Thr	
				160					165					170		
ccc	cgg	att	gag	cgg	gca	ggg	atg	gat	ggc	agc	acc	cgg	aag	atc	att	637
Pro	Arg	Ile	Glu	Arg	Ala	Gly	Met	Asp	Gly	Ser	Thr	Arg	Lys	Ile	Ile	
				175					180					185		
gtg	gac	tgc	gac	att	tac	tgg	ccc	aat	gga	ctg	acc	atc	gac	ctg	gag	685
Val	Asp	Ser	Asp	Ile	Tyr	Trp	Pro	Asn	Gly	Leu	Thr	Ile	Asp	Leu	Glu	
190							195				200				205	
gag	cag	aag	ctc	tac	tgg	gct	gac	gcc	aag	ctc	agc	ttc	atc	cac	cgt	733
Glu	Gln	Lys	Leu	Tyr	Trp	Ala	Asp	Ala	Lys	Leu	Ser	Phe	Ile	His	Arg	
				210					215						220	
gcc	aac	ctg	gac	ggc	tgc	ttc	cgg	cag	aag	gtg	gtg	gag	ggc	agc	ctg	781
Ala	Asn	Leu	Asp	Gly	Ser	Phe	Arg	Gln	Lys	Val	Val	Glu	Gly	Ser	Leu	
				225					230					235		
acg	cac	ccc	ttc	gcc	ctg	acg	ctc	tcc	ggg	gac	act	ctg	tac	tgg	aca	829
Thr	His	Pro	Phe	Ala	Leu	Thr	Leu	Ser	Gly	Asp	Thr	Leu	Tyr	Trp	Thr	
				240					245					250		
gac	tgg	cag	acc	cgc	tcc	atc	cat	gcc	tgc	aac	aag	cgc	act	ggg	ggg	877
Asp	Trp	Gln	Thr	Arg	Ser	Ile	His	Ala	Cys	Asn	Lys	Arg	Thr	Gly	Gly	
				255					260					265		
aag	agg	aag	gag	atc	ctg	agt	gcc	ctc	tac	tca	ccc	atg	gac	atc	cag	925
Lys	Arg	Lys	Glu	Ile	Leu	Ser	Ala	Leu	Tyr	Ser	Pro	Met	Asp	Ile	Gln	
270							275				280				285	
gtg	ctg	agc	cag	gag	cgg	cag	cct	ttc	ttc	cac	act	cgc	tgt	gag	gag	973
Val	Leu	Ser	Gln	Glu	Arg	Gln	Pro	Phe	Phe	His	Thr	Arg	Cys	Glu	Glu	
				290						295					300	
gac	aat	ggc	ggc	tgc	tcc	cac	ctg	tgc	ctg	ctg	tcc	cca	agc	gag	cct	1021
Asp	Asn	Gly	Gly	Cys	Ser	His	Leu	Cys	Leu	Leu	Ser	Pro	Ser	Glu	Pro	
				305					310					315		
ttc	tac	aca	tgc	gcc	tgc	ccc	acg	ggt	gtg	cag	ctg	cag	gac	aac	ggc	1069
Phe	Tyr	Thr	Cys	Ala	Cys	Pro	Thr	Gly	Val	Gln	Leu	Gln	Asp	Asn	Gly	
				320					325					330		
agg	acg	tgt	aag	gca	gga	gcc	gag	gag	gtg	ctg	ctg	ctg	gcc	cgg	cgg	1117
Arg	Thr	Cys	Lys	Ala	Gly	Ala	Glu	Glu	Val	Leu	Leu	Leu	Ala	Arg	Arg	
				335					340					345		
acg	gac	cta	cgg	agg	atc	tgc	ctg	gac	acg	ccg	gac	ttc	acc	gac	atc	1165
Thr	Asp	Leu	Arg	Arg	Ile	Ser	Leu	Asp	Thr	Pro	Asp	Phe	Thr	Asp	Ile	
350							355				360				365	
gtg	ctg	cag	gtg	gac	gac	atc	cgg	cac	gcc	att	gcc	atc	gac	tac	gac	1213
Val	Leu	Gln	Val	Asp	Asp	Ile	Arg	His	Ala	Ile	Ala	Ile	Asp	Tyr	Asp	

	370	375	380	
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc				1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile				
	385	390	395	
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc				1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr				
	400	405	410	
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac				1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn				
	415	420	425	
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc				1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu				
	430	435	440	445
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc				1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro				
	450	455	460	
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac				1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp				
	465	470	475	
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag				1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu				
	480	485	490	
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc				1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala				
	495	500	505	
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac				1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp				
	510	515	520	525
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg				1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu				
	530	535	540	
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc				1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe				
	545	550	555	
atc tac tgg act gac tgg cag cgc agc atc gag cgg gtg cac aag				1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys				
	560	565	570	
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg				1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met				
	575	580	585	
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt				1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys				
	590	595	600	605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac				1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His				
	610	615	620	
gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg				1981
Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met				
	625	630	635	
aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc				2029
Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala				
	640	645	650	
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc				2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile				

655	660	665	
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc			2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser			
670	675	680	685
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc			2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg			
	690	695	700
gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt			2221
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu			
	705	710	715
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac			2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr			
	720	725	730
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg			2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly			
	735	740	745
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg			2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser			
	750	755	760
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc			2413
Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly			
	770	775	780
ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg			2461
Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met			
	785	790	795
acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac			2509
Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr			
	800	805	810
gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag			2557
Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu			
	815	820	825
tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc			2605
Ser Ser Asn Met Leu Gly Gln Glu Arg Val Ile Ala Asp Asp Leu			
	830	835	840
ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca			2653
Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr			
	850	855	860
gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg			2701
Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg			
	865	870	875
aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg			2749
Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu			
	880	885	890
gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac			2797
Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn			
	895	900	905
aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc			2845
Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg			
	910	915	920
tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc			2893
Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys			
	930	935	940
agc ccg ccc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt			2941
Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser			

	945		950		955		
cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg							2989
Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu							
	960		965		970		
cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag							3037
His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys							
	975		980		985		
ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac							3085
Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp							
	990		995		1000		1005
gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac							3133
Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn							
	1010		1015		1020		
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca							3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr							
	1025		1030		1035		
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg							3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu							
	1040		1045		1050		
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc							3277
Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro							
	1055		1060		1065		
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac							3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn							
	1070		1075		1080		1085
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc							3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr							
	1090		1095		1100		
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg							3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu							
	1105		1110		1115		
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg							3469
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu							
	1120		1125		1130		
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg							3517
Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu							
	1135		1140		1145		
gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag							3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys							
	1150		1155		1160		1165
cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag							3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu							
	1170		1175		1180		
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac							3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His							
	1185		1190		1195		
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca							3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser							
	1200		1205		1210		
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att							3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile							
	1215		1220		1225		
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg							3805
Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val							

1230	1235	1240	1245	
ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg				3853
Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro				
	1250	1255	1260	
gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc				3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala				
	1265	1270	1275	
tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag				3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu				
	1280	1285	1290	
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag				3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln				
	1295	1300	1305	
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac				4045
Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp				
1310	1315	1320	1325	
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc				4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe				
	1330	1335	1340	
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc				4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser				
	1345	1350	1355	
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc				4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr				
	1360	1365	1370	
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc				4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro				
	1375	1380	1385	
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt				4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe				
1390	1395	1400	1405	
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc				4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro				
	1410	1415	1420	
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc				4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe				
	1425	1430	1435	
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc				4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys				
	1440	1445	1450	
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg				4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly				
	1455	1460	1465	
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc				4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser				
1470	1475	1480	1485	
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg				4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro				
	1490	1495	1500	
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac				4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr				
	1505	1510	1515	
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att				4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile				

```

1520          1525          1530
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac      4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp
1535          1540          1545
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat      4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp
1550          1555          1560          1565
ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc      4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser
1570          1575          1580
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag      4861
Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
1585          1590          1595
agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac      4909
Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
1600          1605          1610
tca tcc tgacctcggc cgggccactc tggcttctct gtgcccctgt aaatagtttt      4965
Ser Ser
1615
aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt      5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa ctttgtacag      5085
tggagaaata tttataaaact taattttgta aaaca                                5120

<210> 2
<211> 5120
<212> DNA
<213> Homo sapiens

<400> 2
actaaagcgc cgccgccgcg ccatggagcc cgagtgcgcg cggcgcgggc ccgtccggcc      60
gcccgacaac  atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg      109
              Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu
              1          5          10
ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc      157
Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala
15          20          25
gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg      205
Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu
30          35          40          45
gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc      253
Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly
50          55          60
ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg      301
Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val
65          70          75
tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac      349
Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn
80          85          90
cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct      397
Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser
95          100          105
ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg      445
Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr
110          115          120          125

```

gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc	493
Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser	
130 135 140	
cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc	541
Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala	
145 150 155	
ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg gtt gag acg	589
Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr	
160 165 170	
ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att	637
Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile	
175 180 185	
gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag	685
Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu	
190 195 200 205	
gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt	733
Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg	
210 215 220	
gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg	781
Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu	
225 230 235	
acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca	829
Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr	
240 245 250	
gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg	877
Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly	
255 260 265	
aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag	925
Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln	
270 275 280 285	
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag	973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu	
290 295 300	
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct	1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro	
305 310 315	
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc	1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly	
320 325 330	
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg	1117
Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg	
335 340 345	
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc	1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile	
350 355 360 365	
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac	1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp	
370 375 380	
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc	1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile	
385 390 395	
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc	1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr	
400 405 410	

gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac	1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn	
415 420 425	
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc	1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu	
430 435 440 445	
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc	1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro	
450 455 460	
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac	1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp	
465 470 475	
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag	1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu	
480 485 490	
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc	1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala	
495 500 505	
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac	1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp	
510 515 520 525	
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg	1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu	
530 535 540	
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc	1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe	
545 550 555	
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag	1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys	
560 565 570	
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg	1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met	
575 580 585	
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt	1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys	
590 595 600 605	
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac	1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His	
610 615 620	
gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg	1981
Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met	
625 630 635	
aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc	2029
Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala	
640 645 650	
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc	2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile	
655 660 665	
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc	2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser	
670 675 680 685	
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc	2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg	
690 695 700	

gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt	2221
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu	
705 710 715	
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac	2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr	
720 725 730	
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg	2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly	
735 740 745	
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg	2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser	
750 755 760 765	
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc	2413
Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly	
770 775 780	
ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg	2461
Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met	
785 790 795	
acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac	2509
Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr	
800 805 810	
gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag	2557
Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu	
815 820 825	
tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc	2605
Ser Ser Asn Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu	
830 835 840 845	
ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca	2653
Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr	
850 855 860	
gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg	2701
Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg	
865 870 875	
aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg	2749
Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu	
880 885 890	
gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac	2797
Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn	
895 900 905	
aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc	2845
Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg	
910 915 920 925	
tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc	2893
Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys	
930 935 940	
agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt	2941
Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser	
945 950 955	
cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg	2989
Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu	
960 965 970	
cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag	3037
His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys	
975 980 985	

ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac	3085
Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp	
990 995 1000 1005	
gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac	3133
Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn	
1010 1015 1020	
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca	3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr	
1025 1030 1035	
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg	3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu	
1040 1045 1050	
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc	3277
Ser Gly Glu Ala Met Gly Val Leu Arg Gly Asp Arg Asp Lys Pro	
1055 1060 1065	
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac	3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn	
1070 1075 1080 1085	
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc	3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr	
1090 1095 1100	
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg	3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu	
1105 1110 1115	
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg	3469
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu	
1120 1125 1130	
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg	3517
Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu	
1135 1140 1145	
gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag	3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys	
1150 1155 1160 1165	
cat ctc tac tgg atc gac cgc cag cag cag atc atc gag cgt gtg gag	3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu	
1170 1175 1180	
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac	3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His	
1185 1190 1195	
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca	3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser	
1200 1205 1210	
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att	3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile	
1215 1220 1225	
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg	3805
Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val	
1230 1235 1240 1245	
ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg	3853
Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro	
1250 1255 1260	
gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc	3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala	
1265 1270 1275	

tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag	3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu	
1280 1285 1290	
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag	3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln	
1295 1300 1305	
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac	4045
Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp	
1310 1315 1320 1325	
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc	4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe	
1330 1335 1340	
cgg tgt gcg agc ggc cag tgt gtc ctc aaa cag cag tgc gac tcc	4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser	
1345 1350 1355	
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc	4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr	
1360 1365 1370	
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc	4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro	
1375 1380 1385	
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt	4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe	
1390 1395 1400 1405	
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc	4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro	
1410 1415 1420	
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc	4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe	
1425 1430 1435	
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc	4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys	
1440 1445 1450	
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg	4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly	
1455 1460 1465	
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc	4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser	
1470 1475 1480 1485	
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg	4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro	
1490 1495 1500	
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac	4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr	
1505 1510 1515	
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att	4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile	
1520 1525 1530	
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac	4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp	
1535 1540 1545	
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat	4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp	
1550 1555 1560 1565	

```

ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc      4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser
      1570      1575      1580
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag      4861
Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
      1585      1590      1595
agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac      4909
Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
      1600      1605      1610
tca tcc tgacctcggc cgggccactc tggtttctct gtgcccctgt aaatagtttt      4965
Ser Ser
      1615
aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt      5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa ctttgtacag      5085
tggagaaata tttataaact taattttgta aaaca      5120

```

<210> 3

<211> 1615

<212> PRT

<213> Homo sapiens

<400> 3

```

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1          5          10          15
Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
      20      25      30
Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
      35      40      45
Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
      50      55      60
Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
      65      70      75      80
Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
      85      90      95
Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
      100      105      110
Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
      115      120      125
Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
      130      135      140
Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
      145      150      155      160
Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
      165      170      175
Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
      180      185      190
Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
      195      200      205
Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
      210      215      220
Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
      225      230      235      240
Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
      245      250      255

```

-14-

Ile	Tyr	Trp	Thr	Asp	Val	Ser	Leu	Lys	Asn	Ile	Ser	Arg	Ala	Phe	Met
690						695				700					
Asn	Gly	Ser	Ser	Val	Glu	His	Val	Val	Glu	Phe	Gly	Leu	Asp	Tyr	Pro
705					710				715						720
Glu	Gly	Met	Ala	Val	Asp	Trp	Met	Gly	Lys	Asn	Leu	Tyr	Trp	Ala	Asp
				725				730						735	
Thr	Gly	Thr	Asn	Arg	Ile	Glu	Val	Ala	Arg	Leu	Asp	Gly	Gln	Phe	Arg
			740					745					750		
Gln	Val	Leu	Val	Trp	Arg	Asp	Leu	Asp	Asn	Pro	Arg	Ser	Leu	Ala	Leu
		755					760					765			
Asp	Pro	Thr	Lys	Gly	Tyr	Ile	Tyr	Trp	Thr	Glu	Trp	Gly	Gly	Lys	Pro
	770				775					780					
Arg	Ile	Val	Arg	Ala	Phe	Met	Asp	Gly	Thr	Asn	Cys	Met	Thr	Leu	Val
785					790					795					800
Asp	Lys	Val	Gly	Arg	Ala	Asn	Asp	Leu	Thr	Ile	Asp	Tyr	Ala	Asp	Gln
				805					810					815	
Arg	Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Met	Ile	Glu	Ser	Ser	Asn
			820					825					830		
Met	Leu	Gly	Gln	Glu	Arg	Val	Val	Ile	Ala	Asp	Asp	Leu	Pro	His	Pro
		835					840					845			
Phe	Gly	Leu	Thr	Gln	Tyr	Ser	Asp	Tyr	Ile	Tyr	Trp	Thr	Asp	Trp	Asn
	850					855					860				
Leu	His	Ser	Ile	Glu	Arg	Ala	Asp	Lys	Thr	Ser	Gly	Arg	Asn	Arg	Thr
865					870					875					880
Leu	Ile	Gln	Gly	His	Leu	Asp	Phe	Val	Met	Asp	Ile	Leu	Val	Phe	His
				885					890					895	
Ser	Ser	Arg	Gln	Asp	Gly	Leu	Asn	Asp	Cys	Met	His	Asn	Asn	Gly	Gln
			900					905					910		
Cys	Gly	Gln	Leu	Cys	Leu	Ala	Ile	Pro	Gly	Gly	His	Arg	Cys	Gly	Cys
		915					920					925			
Ala	Ser	His	Tyr	Thr	Leu	Asp	Pro	Ser	Ser	Arg	Asn	Cys	Ser	Pro	Pro
	930					935					940				
Thr	Thr	Phe	Leu	Leu	Phe	Ser	Gln	Lys	Ser	Ala	Ile	Ser	Arg	Met	Ile
945					950					955					960
Pro	Asp	Asp	Gln	His	Ser	Pro	Asp	Leu	Ile	Leu	Pro	Leu	His	Gly	Leu
				965					970					975	
Arg	Asn	Val	Lys	Ala	Ile	Asp	Tyr	Asp	Pro	Leu	Asp	Lys	Phe	Ile	Tyr
			980					985					990		
Trp	Val	Asp	Gly	Arg	Gln	Asn	Ile	Lys	Arg	Ala	Lys	Asp	Asp	Gly	Thr
		995					1000						1005		
Gln	Pro	Phe	Val	Leu	Thr	Ser	Leu	Ser	Gln	Gly	Gln	Asn	Pro	Asp	Arg
	1010						1015					1020			
Gln	Pro	His	Asp	Leu	Ser	Ile	Asp	Ile	Tyr	Ser	Arg	Thr	Leu	Phe	Trp
1025					1030					1035					1040
Thr	Cys	Glu	Ala	Thr	Asn	Thr	Ile	Asn	Val	His	Arg	Leu	Ser	Gly	Glu
				1045					1050					1055	
Ala	Met	Gly	Val	Val	Leu	Arg	Gly	Asp	Arg	Asp	Lys	Pro	Arg	Ala	Ile
			1060					1065					1070		
Val	Val	Asn	Ala	Glu	Arg	Gly	Tyr	Leu	Tyr	Phe	Thr	Asn	Met	Gln	Asp
		1075					1080						1085		
Arg	Ala	Ala	Lys	Ile	Glu	Arg	Ala	Ala	Leu	Asp	Gly	Thr	Glu	Arg	Glu
	1090					1095					1100				
Val	Leu	Phe	Thr	Thr	Gly	Leu	Ile	Arg	Pro	Val	Ala	Leu	Val	Val	Asp
1105					1110					1115					1120

Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile	1125	1130	1135
Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala	1140	1145	1150
Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr	1155	1160	1165
Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr	1170	1175	1180
Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly	1185	1190	1195
Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro	1205	1210	1215
Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly	1220	1225	1230
Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln	1235	1240	1245
Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe	1250	1255	1260
Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys	1265	1270	1275
Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro	1285	1290	1295
Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp	1300	1305	1310
Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp	1315	1320	1325
Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala	1330	1335	1340
Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp	1345	1350	1355
Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro	1365	1370	1375
Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly	1380	1385	1390
Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln	1395	1400	1405
Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His	1410	1415	1420
Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro	1425	1430	1435
Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser	1445	1450	1455
Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu	1460	1465	1470
Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser	1475	1480	1485
Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro	1490	1495	1500
Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn	1505	1510	1515
Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met	1525	1530	1535
Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr	1540	1545	1550

Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
 1555 1560 1565
 Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
 1570 1575 1580
 Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
 1585 1590 1595 1600
 Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610 1615

<210> 4
 <211> 1615
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ser
 20 25 30
 Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45
 Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60
 Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80
 Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95
 Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110
 Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125
 Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
 130 135 140
 Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
 145 150 155 160
 Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr Pro Arg Ile
 165 170 175
 Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
 180 185 190
 Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
 195 200 205
 Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
 210 215 220
 Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
 225 230 235 240
 Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
 245 250 255
 Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
 260 265 270
 Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
 275 280 285
 Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
 290 295 300

Gly	Trp	Ser	His	Leu	Cys	Leu	Leu	Ser	Pro	Ser	Glu	Pro	Phe	Tyr	Thr	
305					310					315					320	
Cys	Ala	Cys	Pro	Thr	Gly	Val	Gln	Met	Gln	Asp	Asn	Gly	Arg	Thr	Cys	
				325					330					335		
Lys	Ala	Gly	Ala	Glu	Glu	Val	Leu	Leu	Leu	Ala	Arg	Arg	Thr	Asp	Leu	
			340					345					350			
Arg	Arg	Ile	Ser	Leu	Asp	Thr	Pro	Asp	Phe	Thr	Asp	Ile	Val	Leu	Gln	
		355				360					365					
Val	Asp	Asp	Ile	Arg	His	Ala	Ile	Ala	Ile	Asp	Tyr	Asp	Pro	Leu	Glu	
	370				375					380						
Gly	Tyr	Val	Tyr	Trp	Thr	Asp	Asp	Glu	Val	Arg	Ala	Ile	Arg	Arg	Ala	
385					390					395					400	
Tyr	Leu	Asp	Gly	Ser	Gly	Ala	Gln	Thr	Leu	Val	Asn	Thr	Glu	Ile	Asn	
			405						410					415		
Asp	Pro	Asp	Gly	Ile	Ala	Val	Asp	Trp	Val	Ala	Arg	Asn	Leu	Tyr	Trp	
		420						425					430			
Thr	Asp	Thr	Gly	Thr	Asp	Arg	Ile	Glu	Val	Thr	Arg	Leu	Asn	Gly	Thr	
		435				440						445				
Ser	Arg	Lys	Ile	Leu	Val	Ser	Glu	Asp	Leu	Asp	Glu	Pro	Arg	Ala	Ile	
	450				455					460						
Ala	Leu	His	Pro	Val	Met	Gly	Leu	Met	Tyr	Trp	Thr	Asp	Trp	Gly	Glu	
465					470					475					480	
Asn	Pro	Lys	Ile	Glu	Cys	Ala	Asn	Leu	Asp	Gly	Gln	Glu	Arg	Arg	Val	
			485						490					495		
Leu	Val	Asn	Ala	Ser	Leu	Gly	Trp	Pro	Asn	Gly	Leu	Ala	Leu	Asp	Leu	
		500						505					510			
Gln	Glu	Gly	Lys	Leu	Tyr	Trp	Gly	Asp	Ala	Lys	Thr	Asp	Lys	Ile	Glu	
		515					520					525				
Val	Ile	Asn	Val	Asp	Gly	Thr	Lys	Arg	Arg	Thr	Leu	Leu	Glu	Asp	Lys	
	530					535					540					
Leu	Pro	His	Ile	Phe	Gly	Phe	Thr	Leu	Leu	Gly	Asp	Phe	Ile	Tyr	Trp	
545					550					555					560	
Thr	Asp	Trp	Gln	Arg	Arg	Ser	Ile	Glu	Arg	Val	His	Lys	Val	Lys	Ala	
			565						570					575		
Ser	Arg	Asp	Val	Ile	Ile	Asp	Gln	Leu	Pro	Asp	Leu	Met	Gly	Leu	Lys	
		580						585					590			
Ala	Val	Asn	Val	Ala	Lys	Val	Val	Gly	Thr	Asn	Pro	Cys	Ala	Asp	Arg	
		595						600				605				
Asn	Gly	Gly	Cys	Ser	His	Leu	Cys	Phe	Phe	Thr	Pro	His	Ala	Thr	Arg	
	610					615					620					
Cys	Gly	Cys	Pro	Ile	Gly	Leu	Glu	Leu	Leu	Ser	Asp	Met	Lys	Thr	Cys	
625					630					635					640	
Ile	Val	Pro	Glu	Ala	Phe	Leu	Val	Phe	Thr	Ser	Arg	Ala	Ala	Ile	His	
			645						650					655		
Arg	Ile	Ser	Leu	Glu	Thr	Asn	Asn	Asn	Asp	Val	Ala	Ile	Pro	Leu	Thr	
		660						665					670			
Gly	Val	Lys	Glu	Ala	Ser	Ala	Leu	Asp	Phe	Asp	Val	Ser	Asn	Asn	His	
		675						680				685				
Ile	Tyr	Trp	Thr	Asp	Val	Ser	Leu	Lys	Asn	Ile	Ser	Arg	Ala	Phe	Met	
	690					695					700					
Asn	Gly	Ser	Ser	Val	Glu	His	Val	Val	Glu	Phe	Gly	Leu	Asp	Tyr	Pro	
705					710					715					720	
Glu	Gly	Met	Ala	Val	Asp	Trp	Met	Gly	Lys	Asn	Leu	Tyr	Trp	Ala	Asp	
			725						730					735		

Thr	Gly	Thr	Asn	Arg	Ile	Glu	Val	Ala	Arg	Leu	Asp	Gly	Gln	Phe	Arg		
			740					745					750				
Gln	Val	Leu	Val	Trp	Arg	Asp	Leu	Asp	Asn	Pro	Arg	Ser	Leu	Ala	Leu		
			755				760					765					
Asp	Pro	Thr	Lys	Gly	Tyr	Ile	Tyr	Trp	Thr	Glu	Trp	Gly	Gly	Lys	Pro		
			770			775					780						
Arg	Ile	Val	Arg	Ala	Phe	Met	Asp	Gly	Thr	Asn	Cys	Met	Thr	Leu	Val		
785					790					795					800		
Asp	Lys	Val	Gly	Arg	Ala	Asn	Asp	Leu	Thr	Ile	Asp	Tyr	Ala	Asp	Gln		
				805					810					815			
Arg	Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Met	Ile	Glu	Ser	Ser	Asn		
			820					825					830				
Met	Leu	Gly	Gln	Glu	Arg	Val	Val	Ile	Ala	Asp	Asp	Leu	Pro	His	Pro		
			835				840					845					
Phe	Gly	Leu	Thr	Gln	Tyr	Ser	Asp	Tyr	Ile	Tyr	Trp	Thr	Asp	Trp	Asn		
850						855					860						
Leu	His	Ser	Ile	Glu	Arg	Ala	Asp	Lys	Thr	Ser	Gly	Arg	Asn	Arg	Thr		
865					870					875					880		
Leu	Ile	Gln	Gly	His	Leu	Asp	Phe	Val	Met	Asp	Ile	Leu	Val	Phe	His		
				885					890					895			
Ser	Ser	Arg	Gln	Asp	Gly	Leu	Asn	Asp	Cys	Met	His	Asn	Asn	Gly	Gln		
			900					905					910				
Cys	Gly	Gln	Leu	Cys	Leu	Ala	Ile	Pro	Gly	Gly	His	Arg	Cys	Gly	Cys		
			915				920					925					
Ala	Ser	His	Tyr	Thr	Leu	Asp	Pro	Ser	Ser	Arg	Asn	Cys	Ser	Pro	Pro		
			930			935					940						
Thr	Thr	Phe	Leu	Leu	Phe	Ser	Gln	Lys	Ser	Ala	Ile	Ser	Arg	Met	Ile		
945					950					955					960		
Pro	Asp	Asp	Gln	His	Ser	Pro	Asp	Leu	Ile	Leu	Pro	Leu	His	Gly	Leu		
				965					970					975			
Arg	Asn	Val	Lys	Ala	Ile	Asp	Tyr	Asp	Pro	Leu	Asp	Lys	Phe	Ile	Tyr		
			980					985					990				
Trp	Val	Asp	Gly	Arg	Gln	Asn	Ile	Lys	Arg	Ala	Lys	Asp	Asp	Gly	Thr		
			995			1000						1005					
Gln	Pro	Phe	Val	Leu	Thr	Ser	Leu	Ser	Gln	Gly	Gln	Asn	Pro	Asp	Arg		
			1010			1015					1020						
Gln	Pro	His	Asp	Leu	Ser	Ile	Asp	Ile	Tyr	Ser	Arg	Thr	Leu	Phe	Trp		
1025					1030					1035					1040		
Thr	Cys	Glu	Ala	Thr	Asn	Thr	Ile	Asn	Val	His	Arg	Leu	Ser	Gly	Glu		
				1045					1050					1055			
Ala	Met	Gly	Val	Val	Leu	Arg	Gly	Asp	Arg	Asp	Lys	Pro	Arg	Ala	Ile		
			1060					1065					1070				
Val	Val	Asn	Ala	Glu	Arg	Gly	Tyr	Leu	Tyr	Phe	Thr	Asn	Met	Gln	Asp		
			1075				1080					1085					
Arg	Ala	Ala	Lys	Ile	Glu	Arg	Ala	Ala	Leu	Asp	Gly	Thr	Glu	Arg	Glu		
			1090			1095					1100						
Val	Leu	Phe	Thr	Thr	Gly	Leu	Ile	Arg	Pro	Val	Ala	Leu	Val	Val	Asp		
1105					1110					1115					1120		
Asn	Thr	Leu	Gly	Lys	Leu	Phe	Trp	Val	Asp	Ala	Asp	Leu	Lys	Arg	Ile		
				1125					1130					1135			
Glu	Ser	Cys	Asp	Leu	Ser	Gly	Ala	Asn	Arg	Leu	Thr	Leu	Glu	Asp	Ala		
			1140					1145				1150					
Asn	Ile	Val	Gln	Pro	Leu	Gly	Leu	Thr	Ile	Leu	Gly	Lys	His	Leu	Tyr		
			1155				1160					1165					

Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
 1170 1175 1180
 Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
 1185 1190 1195 1200
 Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
 1205 1210 1215
 Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
 1220 1225 1230
 Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
 1235 1240 1245
 Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
 1250 1255 1260
 Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
 1265 1270 1275 1280
 Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
 1285 1290 1295
 Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
 1300 1305 1310
 Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
 1315 1320 1325
 Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
 1330 1335 1340
 Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
 1345 1350 1355 1360
 Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
 1365 1370 1375
 Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
 1380 1385 1390
 Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
 1395 1400 1405
 Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
 1410 1415 1420
 Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
 1425 1430 1435 1440
 Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
 1445 1450 1455
 Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
 1460 1465 1470
 Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser
 1475 1480 1485
 Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
 1490 1495 1500
 Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
 1505 1510 1515 1520
 Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
 1525 1530 1535
 Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
 1540 1545 1550
 Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
 1555 1560 1565
 Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
 1570 1575 1580
 Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
 1585 1590 1595 1600

Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1605 1610 1615

<210> 5

<211> 3096

<212> DNA

<213> Homo sapiens

<400> 5

catcttctca	cacgatctct	cgcttcgcac	tccttccttt	gattgggttt	caccattttac	60
tcagacgacg	gtccttcttc	gatcttttga	cattcttcta	tcactacta	ccttcataacc	120
cagctccgct	ccetaatatt	catgcgcgga	tggcccattc	cgtggtgaaa	attcccttct	180
actctgctaa	tctgctgttc	tctctccctc	ccgtcgggtt	ctgctcctgc	cacgttctcc	240
cctctcccca	ccaaaggctg	ggttttcttt	gtcagggctc	ctttcccctt	tggaagaagg	300
ggggctgtat	ggccttggtg	cgaggccctc	cagtacagag	atcccccatc	accagagatt	360
ccacaggccc	tggtagggag	gagggggagc	agaagaggag	gtgccatctt	tgctgtctgg	420
ggaaggccag	gggccacca	cacagagctc	tcccatttgc	tgtggaccct	ggggccactg	480
cccagttcct	tccaaaggaa	agccagctcc	ccaggtgggtg	ggagagtgat	atggcttcct	540
cttaaaactta	gggaattgag	tgtgtggttg	cttctaagtg	ccttagaagc	cgggagcggc	600
tcctggaaag	agcctgcctg	ccacagcggg	ccttaccctg	gctgtgcca	cagatgtccc	660
tggggcctgc	cgctcctgcc	cggctctcct	ggcctcccc	ggtgtgggtt	gggaaaagca	720
cagcaaatta	aaaaacacct	ccatctctgg	cctttgaaga	atgcatctga	acagccgaga	780
gtgtaaaccg	tggtagaatg	tggctcttcc	agtttgggga	gaagcagggc	agagctgggg	840
cttttgtacc	cagggtttcc	aagagctcct	gcctccctcg	gctgggctgg	ccagggcccc	900
ccgctgggac	ctccagctgt	aatagggaag	gttttactgg	gttgctggcc	actgtggact	960
gcccctaagg	gcaggtatgc	ctgcctttac	ccgggttccc	ctcctgcctg	gaagatacag	1020
cccatgggag	gcctgttgtc	tgtgggatcc	tccagcatca	gagacactgg	ggccagcgtc	1080
tgcttggtga	ggtgcaggcc	tggcaggccc	ggtccccac	ctgcttgagc	accacgggtg	1140
gtgggggctc	gctgcctccc	gagacaatct	atgtcattgt	tgtccaagga	agctaattta	1200
gagtagaaag	ttcctgttcc	agtcccactc	tgtgcgtgtg	ttagcagggg	actctcgggc	1260
cggagctggg	tccaccctgg	tagggggact	tcattggggc	tgggcgacag	cactgtgtat	1320
ttgtgtgtgt	gtgtgtttgt	gtgtgtgtgt	gtctgaggag	gtggaccagt	ttctcaaaag	1380
gctgtgacc	ccaagaacca	aggaatttca	gcctgggtgg	atcacacctt	cactggtgag	1440
tgggacaagc	tgggggccct	cgccacagga	gcagccaggg	catggggcac	agtggcctc	1500
attcacaaaa	tgggagtata	agtgatccct	gctctggcgg	ccaggacgat	gagtgggaac	1560
acaccgtgtg	ggggctgcct	ggcctgggtg	tgcgcggggt	gtccttgttg	gtgatggttc	1620
cacctctctg	tggcaccagt	gccctctggg	tctcacacac	aactctcttc	ccagcgaagg	1680
cccctcctgc	cctcaggcct	cagtgtctgt	tccgtctcgg	aaggccccag	gagctcctgc	1740
atcctgggcg	tgattcctgt	gtgcctgcag	acccctcgc	ggctgccatc	tcactctttg	1800
gtgcacctgt	tggccagacc	tcctggtagc	gggtgtgtga	ctcccctgaa	tgtgccgggg	1860
cctgggggca	gggacctggg	ctcctccctc	actgagtggg	gggaactcag	tgtcttggag	1920
ttgggggtgcc	tgcaggctgg	gtggtgcagg	tgaaatgcag	acctctcagc	tgggtgtcca	1980
gagcagctgc	cttccccgc	ccgagggact	tcaccgcag	cccagtcagg	ggtggcgcct	2040
gggtgcatcg	ccgcaggct	gggtaggggt	ggagcctggg	tggccctgcc	tgtgagctgc	2100
atagttgtcg	cctttgaccc	tgagttttct	tcgttatctg	tttggacctg	tttggggcag	2160
gcaggggatg	agatctgaag	ataaatgcct	tagctgtgac	catctccttt	tgtgagaggt	2220
caatgtccag	ttccgctgca	gttataacat	cccatttttt	gatttctttt	tattttttcc	2280
tttttctttt	tgagatggag	tctcgtctctg	tcaccagggc	tggagtgcaa	tggggtgacc	2340
tcagctcact	gcaacctcca	cttctcgggt	tcaagtgatt	ctcctgcctc	agcctcctga	2400
ctagcagggg	ttacaggcgt	gagccaccac	gccagctaa	tttttgtatt	tttagtagag	2460
gcaaggtttc	gtcatgttgg	ccaggtgtgt	ctcaaactcc	tggccttaag	tgatctgcc	2520
gcctcggcct	cccaaagtgc	tgagatgaca	ggtgtgagcc	accgtgcccc	gcccagaact	2580
ctttaattcc	cacctgaaac	ttgccgcctt	aagcaggtcc	ccagtctccc	tcccctagtc	2640

cctgggtccca	ccattctgct	ttctgtctca	atgaatttgc	ctaccgtaag	tacctcatat	2700
aaattgaatc	ataaagtatt	tgtcttttta	tatctggctt	atttcactta	gcataacatt	2760
cttaagtttc	atccatgttg	tagcatgtgt	cagaatctct	ctcttttttt	tttttttttt	2820
tttttttttt	ttttgcagac	agagtctcgc	tctgtcatct	agactggagt	tcagtggcac	2880
gatctcgggt	cactgcaaca	tctgcctcct	gggtccaagc	aattctcctg	cctcagcctc	2940
cttagcagct	ggaactacag	gcgcgtgcc	ccatgccttg	ctaatttttg	tatttttatg	3000
tggaggcagg	gtttcaccat	cttggccagg	ctgggtctcga	attcctgggc	ttcaccacgg	3060
gggcccgaag	gacccgggca	aagcgtggag	gggagg			3096

<210> 6

<211> 26928

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (12044), (12489), (26433), (26434), (26435), (26436), (26439), (26441)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 6

gaagaccaag	ggcacacagc	gaggcagttt	cagggcgggc	agcctggggc	cccacggggc	60
ggccccggac	acttgttctc	acctgtggag	ggcagagaag	ggaacagggg	gagaagtggc	120
cggtcgggag	tggaggtggg	tttgaggttt	tactgtaaac	taaatgtgta	ccctctacct	180
tagttatgaa	ttatgagaca	cgaagactgc	gaaacagaca	cactcctcta	aaagtgcctc	240
taggctgaca	gggagaaagt	cccgccaggc	tcccagacgc	cacctttgag	tccttcaaca	300
agcccgccag	ggcctcttgc	ccaccggtgt	cagctcagcc	actgaaccct	ccaggaagaa	360
gacgtgctgg	taggagaaga	atctcaccca	ggcacagcct	ggaaggggca	cagaaggggc	420
tccggaacca	gcaagcccaa	gttggaaactc	ccagtctgct	actttctaga	acgactgtgc	480
ccttggcggg	tctaagtaga	acctctccgc	gcactctttc	ctccttttga	aagtggggac	540
agcaatggcc	accttgcagg	ttcagagagg	gcttgcagta	cctcacagaa	ctgagtggcc	600
gtgaacgtgt	gtgttcctcc	agatttgtga	cagctttgcc	aggctggagt	caggctgaac	660
gcctctgccc	tcatgggggt	tatattctag	gaagaccaac	aaaaacaaga	agacggaaaa	720
ttaaaacaac	aaaagcccca	ttgacaggcc	gtgaagaatg	ccatgaaaaa	tgaatggcgt	780
tgtgctgcag	tctttggggg	aacgggctta	cggaagaaag	gacacttgag	ctgctaccaa	840
tgagcagccg	tccggtggga	gggcagttca	ggaagagcag	acatccactg	aggaggcgct	900
ggggcagagg	gcagcctggt	cgctggattc	gggggaggaa	ccacatcagg	ccatgagctg	960
gagctggtgg	tagaatgtac	aggagaggcc	agccaggggc	agctcatgtc	agacctcaag	1020
cggggaagat	gaatcgagaa	tgcaccccac	gagcaatggg	aagccagtct	acgatttaag	1080
cagcaaaaaa	attttccctt	cttccaccct	gcattccagct	ctaccagcac	agcctgggggt	1140
tctattttca	agatagaata	gacccagact	cccagctctt	cttacacttc	tactactgcc	1200
acctgtcacc	cactcatgcg	tccccacttg	cagcctcgac	ccccctccac	ctgatctcat	1260
ggcagccagg	gaagctccag	ggctcgtgag	ggctgccatc	tcaggaaaaga	agcaaaagcc	1320
ttcggcacct	gcagggcctg	ctccaaccac	acttcttctc	tgacctctca	gcttcccttag	1380
ccactccctt	cccacatctc	acctgtctcc	agccacagtg	gtgtctctgt	gggttctcaa	1440
acacaccagg	tgcactcctg	cctcaggggc	tttgtgcttg	ctgttctctg	ctgggactct	1500
tttttttttt	tttttttttg	agacagggtc	tcactctgtg	gcccaggctg	gagtgtagtg	1560
gtgtgatcgt	agctcattgc	aacctcaaac	tcttgggctc	aagcaatcct	cccacctcag	1620
cctctcaagt	agttagcttt	tgttgttttg	ttttgagatg	ggatctcact	ctgttgccca	1680
ggctggagtg	cagtggggca	atcttggctc	accacaacct	ctgcctccca	ggctcaagca	1740
attctcctgc	ctcagcctcc	caagtagctg	ggattacagg	catgtgccac	cacgcccagc	1800
ttatttttgt	attttttagta	gagacagggt	ttcaccatgt	tgggtctggct	ggtcttgaac	1860
tcctggcctc	agatgatcca	cctgcctcgg	cctcccaaag	tgctgggatg	acaggcatga	1920
gcctgtctct	agtagttagg	actacagaga	ggggccatca	tgcttgggtg	tcctcccacc	1980
ttttctgctc	caactctttc	accccactta	gcctcgtggc	tcactctctt	acctcttcag	2040

ctcctcagtc	aggcctgagg	acccctgttg	aaaattgcaa	accacacccc	ccaccaccac	2100
caccactat	tgccagcact	ttctactcca	tttctctgct	ttacttttct	cctttgtact	2160
catcaccacc	tgactcatta	catgtttacg	tatctttctt	ctctccacta	gcatggaagc	2220
tccaggagag	cagagagtgt	agttttatct	cctgatgtgt	ttcctgtgcc	cgtaccaggg	2280
cctagcacac	agtaggtgct	cagtaaattgt	gtgttggtatg	aacaaatata	gtgaaaggat	2340
ctgatctaca	tttataaaga	aggcactctg	gctgctgagt	ggggatgaga	ctgtcaggag	2400
gaaagaggcc	cctgtggggg	cctggccagc	aggtgggtac	aatggtagca	gccaggagag	2460
agggcctctt	ggactcaagt	ggatggggcc	tgctcagggc	tccggccaca	ggaacaaagg	2520
gaagggggcc	caggatggcc	tgtcatagag	gacacattac	aactggccca	aagttcaagt	2580
caggtttcta	aatttgggaa	gggatacaga	aaaactaaag	actctactgg	acagtcagtt	2640
attgaaatga	ttacatagaa	aatgtaccaa	gaattaaaaa	aaaaaaaaaa	aagcattatg	2700
aagggggcac	cagagactcc	cagagaggaa	agggactatg	ggctggatgc	ggtgactcac	2760
acctataatc	ccagcacttt	gggaggccga	ggagggtgga	tcacgaggtc	aggagttcaa	2820
aaccagccta	ggcaacatgg	taaaaccccc	gtttctacta	aaaatacaaa	aaattagctg	2880
ggcatggcag	catgtgcctg	taatcccagc	tactcgggag	gctgaggcag	gagagttgct	2940
agaacccagg	aggcagaggt	tgcaagtgagc	cgagattgag	ccactatgct	ccagcttggg	3000
cgacagagca	agactccgtc	tctaaaaaaa	agaaaaaaa	ggccagatga	ggtggctcat	3060
gctgtaatc	ccagcacttt	gggaggccga	ggtgggtgga	tcacgaggtc	aggagatcga	3120
gacctcctg	gctaactagg	tgaactcca	tctctactta	aaatacaaaa	aattagccgg	3180
gcgtggtggc	gggcacctgt	agtcccagct	acttgggagg	ctgaggcagg	agaatggcgt	3240
gaacctggga	ggcgagactt	gcagtgagcc	gagattgcgc	cactgcactc	catccagcct	3300
gggcgacaga	gttagactcc	gtctcaaaaa	aaaaaaaaaa	aaaaaaatta	gctgattagt	3360
tgggcttggg	ggcgggcgcc	tgtaatccca	actactcggg	aggctgaggc	gggagaatca	3420
cttgaacccg	ggaggcagag	gttgcaatga	gccgatatca	cgccactaca	ctccagcctg	3480
ggcgacagag	caagactcca	tctcaaaaaa	gaaaaaaa	aagaaagggg	ctgtgctgtg	3540
gcctgggacc	caaagcacac	tactgcaagg	tcccagggtg	cctgactcca	accggagcct	3600
tgagaacatt	catttgcaaa	gaatgaatta	aaattcagca	ctattttatt	ctgcaggatt	3660
ccagcacccc	aaggacagtc	attttttagac	ccttcagtaa	cgtaataagt	aaccggagga	3720
tgtgctgagc	ttccacttcc	ccagacgggt	gcctgtcaca	gctcatcagg	ccaacaaact	3780
tttcttaggc	ctcaaatttg	gaaatgttca	ctctcagttc	gttccttaga	tgcaagtcca	3840
tcccaatgaa	gtaacagggg	ctcagcacct	gtccaatctc	attgcttccg	gggacagggg	3900
cccatgagga	tgtcgtttca	gcocggtgac	acttgggcaa	agtgcctttt	ggtttccctc	3960
ccaggctgga	acgtgctggc	tctgtgaagt	tacgtcgggc	acaagagccc	cccccaacc	4020
ggcaggactg	actgctgtgg	tcagaggcgc	ccctggggct	ttgggagcca	cagaatcttc	4080
ctgagggcag	cgccggagga	ggccccagtg	agagtgccea	ctgccaggct	cattcctcag	4140
gctgccgcag	gcctctcccc	aaaacaggca	atgcttctca	gcaacctgcc	ccaggagcag	4200
gccagggaag	gccgcocatg	gcctacagtg	ctgggctctg	gagggcctgg	ttggtaacag	4260
gccatggttt	ctatgagcca	gctgggggtg	gaaggacata	ggctggattc	acctctctgg	4320
gcctcagttt	ctgcattcaa	aaagtgggaa	tcatgatata	tgctctattt	cttatctctc	4380
agtgcctgatg	tgaacctoca	ataagacttt	taaaaatact	ctttctacct	tacttttatt	4440
tttcatttat	tttaagataa	tgtctagctg	tctcaccag	gctggagtgc	agtgggtgtga	4500
ttacggctca	ctacagcctt	aacctccag	gctcaagtga	tcctcctacc	acagcctccc	4560
aagtagctgg	aactacaggc	atgcaccacc	gcacctggat	aattttttct	tttgagacaa	4620
ggtttcactc	tgttgcccag	gctggagtgc	agtgggtgcac	tcttggtcca	ctgcagcctc	4680
aacctccctg	ggcttaggtg	atcctcacac	ttcagttctc	caagtagctg	ggactacagg	4740
tatgtgccag	tacaccagc	taatattttt	gaaggatggg	gtttcactat	attgccaggg	4800
ctggctctga	actccagggt	ttaagcaatc	taccttctc	agcctgcaa	agtgctagga	4860
ttataggtat	gagccacccc	ccggcctata	atcctaccac	tttaaaaaag	cctgtaattt	4920
tagcacttta	aaaaattttt	ctaaattttt	tatagagatg	ggggacagct	gtgggtctcac	4980
tgtgttgccc	aggctgggtct	tgaactccta	ggatcaagcc	atcctcctgg	cctggcctcc	5040
caaagtgttg	ggattataag	cataagcctt	accttacctt	ttttttttga	ggtgcagttt	5100
tgttcttggt	gctcaggctg	gagtgcaatg	gcaagatctt	ggctcactgc	aacctccacc	5160
tcccgggttc	aagcaattct	cctgcctcag	cctcccgagt	agctgggatt	acaggcatgc	5220
gccaccacac	ccagctaatt	ttgtattttt	agtagagatg	gggtttctct	atatacctta	5280

attttaaagc	actgcattca	tgtaaattgt	gattaacatg	gattcaagag	agggagttag	5340
gatgaatgag	ccaggcagtc	acctcggtcg	tcaccctcca	cttctctcct	ccttctgaca	5400
gtcatcgctc	atccgtttct	gcagctgttt	gtttgactct	cctgatcatt	ttgcttgcca	5460
cataacttgc	ctcctgggaa	agaatgccct	gggcaggccc	acatgagtag	tgaaaaataa	5520
tctgcagtga	aaaataaaac	taagtagtct	ggccacacaga	gcagtcttat	tttttctactg	5580
cagatgaagg	agttgacatt	caggcttcat	tctcatttat	aagtgtttta	aagacacata	5640
cagtggattg	aacagtggcc	ttcaaaaaga	tgtatctaca	tcctaattccc	tgggacctgt	5700
gaatgttaac	caagttagga	aaagggctct	cccggtgtgc	attaagtttag	agatcttgag	5760
atgaggagct	catcgtggat	tatccagggtg	gacctgcat	ccaaggacaa	atggctcctta	5820
gaaaagaaaa	gcagaggctg	ggcacagtgg	ctcaagcctg	taatcccagc	acttttgagag	5880
gccgaggtgg	gtggatcacc	taaggctcatg	agttcgagag	cagcctggcc	aacatgatga	5940
aatcccctct	ctactaaaaa	tacaaaaatt	agcaaggcat	ggtggcgggt	gcctataatc	6000
ccagctactc	aggaagctga	ggcaggagaa	tggcttgcat	ctgggaggcg	gagggtgcat	6060
tgagccaaga	tgcgcccact	gcactccagc	ctgagggaga	aaagtgaac	tctgtctcat	6120
aaaagaaaaag	aaaagcagac	agagatctga	gacagaagag	gagagtgaag	gaaaaaaggc	6180
catgtgaaga	tgaggcagag	gttggagcca	tgcagccaca	agccaaggaa	tacctggagc	6240
cccagaagtt	gcaagaggta	ggaagaagcc	ctccctagag	cctccagacg	gagcacagcc	6300
ctgccaacac	ctccacctca	gacttctggc	ctccagcact	gtgagataat	caactgctgt	6360
tgttttaagc	caccagattt	gtggtaattt	gttatggcag	ccacaggaaa	ctaatacagt	6420
acctaattct	cacaaaccca	tcttacagaa	aaggaaaactg	aagtcagaga	ggtagtggct	6480
tgtgcagtgt	gttaggccat	tcttgattta	ctataaagaa	atacctgagg	ccgggcatgg	6540
tggctcacgc	ctgtaatccc	agcactttgg	gaggccaagg	tgagtggatc	acttgaggctc	6600
aggagttaa	gaccagcctg	gacaacatgg	tgaaccccca	tttctactga	aaatatgaaa	6660
attagccagg	catggtggcg	tgcactgtga	gtcccagcta	ctcaggaggc	tgaggcagga	6720
gaatcacttg	cgcccgggag	gaggaggttg	tagtgagcca	agattgtgcc	actgcactcc	6780
agcctgggag	acaagagaga	aaccctgtct	caaaataaat	aaaaaacaaa	taaacacctg	6840
agactgggta	gtttataaag	aaaggggtta	actggctccc	ggttctgcag	gctgtacaag	6900
catggtgccg	gcactctgct	ggttgctggg	aaggcttcag	ggagttttac	tcatcgtgga	6960
aggcagagcc	agagcaggtg	catcacacag	caaaagcagg	agcgagagag	agagagagca	7020
gggagggtgtg	cacactttta	aatgagcaga	tctcacgaga	actcaccatt	gcaaggacag	7080
caccaagcca	cgagggtct	gcccccatga	cccaaacctc	ccactaggcc	ccacccccaa	7140
cattgggaat	tacagttcaa	catgaggttt	ggggggacaa	atatccaaac	tatatcattc	7200
caccctggc	ccccagatc	tcatgttctt	ctcacattgc	aaaatatagt	catgccttcc	7260
cagtagcccc	ccaaagtctt	aactcatccc	agcattaact	caaaaatccc	attcccaagt	7320
ccaacgtctc	atctgaagat	gagttccttt	cacctacaag	actgtaaaaa	tgaaaacagt	7380
tatttactgc	tgagatacaa	tgggggcata	ggcattaggt	aaacattcct	gttccaaaag	7440
ggagaaatcg	gtcaaaaaga	aggggctata	ggccccagc	aagtccaaaa	cccagcagag	7500
caatcatatc	atcttaagc	tccaaaataa	cctccttaaa	ctccatgtcc	catagccagg	7560
gcacactggt	gcaaggggca	ggctcccaag	gccttgggca	gctctattcc	tgcggctttg	7620
cagaattcag	tccccatggc	tgctcttaca	gattggagat	gagggcctgc	ggcttttcca	7680
ggtgcagggt	gcaagctgct	ggtgatctac	cattctgggg	tgtggatggt	ggcggccccg	7740
tcccgagct	ccactaggca	ttgtcccagt	ggggactcta	tgtggggcct	ccaacccac	7800
atttccccctc	caatgggaag	gctctgcccc	tgcagcagcc	ttcttctctg	gctcccaggc	7860
tttctcatac	atcctctgac	atctaggtgg	atggtgtcaa	gcttcttca	ctcttgcaact	7920
ctgcacacct	acaggcttaa	caccacatgg	aagctgccaa	ggtgtatggc	tggaaccctc	7980
tgaagcagca	gcctgagctg	tgactatggc	cctttgagcc	aaggctggag	ctggaacagt	8040
ctagatgcag	gcaggagca	gtgtcctgag	gctgtgcaga	gcagcagggc	cctgtgcctg	8100
gacaatgaaa	ccattctttc	ctcctcatcc	tctgggcctg	tgatgggagg	gttgtggaag	8160
atctctgaaa	tgcttttgag	gcctttttgc	ctctgaggcc	tatttcttat	tgtctcagtt	8220
attggcagtc	ggctcctttt	tagttatgca	aatcctctag	caagagggtta	ctccactgcc	8280
ggcttgaact	cctctcctga	aaaagctttt	tctttctttg	tcacatggcc	aggctgcaaa	8340
ttttccaaac	ttttatgctc	tgttttacct	ttaaatataa	cttctaactt	taattcattt	8400
atttgctcct	gcatttgagc	atagggaatt	caaagaagct	gggccacatc	ttgaatgctt	8460
tgctgcttca	aaatttatgg	ccacgcttgg	tggctcacac	ctgtaatccc	agcacttttg	8520

gaggcctagg	tgggcagatc	acgagatcag	gagatcgaga	ccatcctggt	caacatggtg	8580
aaacccatct	ctactaaaaa	tacaaaaaaa	ttagcttggt	gtggtggcgc	agacctgtag	8640
tcccagctac	tggagagggt	gaggcaggag	aattacttga	acctgggagg	cagagggttg	8700
agttagccca	gatcatgcc	ctgcactcca	gcctgggtgac	agaataagat	ttgatctcga	8760
aaggaaggaa	ggaaggagga	agggaagaaa	tgtcttcccc	ccagatgtcc	tgggtcatcc	8820
ctcttatggt	caaacttcaa	cagatcccta	gggcatgaaa	ataatacagc	caaattatct	8880
gctaaggcat	aacgaaagt	acctttgtct	cagttcccaa	taagttcctc	atttccatct	8940
gagactcatc	accctggcct	tggcttggtc	atatcactgt	cagcattttg	gtcacaatca	9000
tttaaccagc	taatcgagg	gctgaggcaa	gaggatcact	tgaacccagg	aggttgaggc	9060
tgcagtgagc	tgtgatcaca	tactgcagt	ccagcttggg	caacagagca	agatcctgtc	9120
tcaataaata	aataaataaa	tacataaata	acttaagttt	atttaaagct	gcattctttg	9180
caccatggag	aaaggccagg	ccagctcctt	ctctctttct	gcacgtgttc	ctcccacctc	9240
agctgcctct	gctcctcaag	gaggaacaga	gggagtagga	aaggccatcc	caggaggccc	9300
agcaccat	gacctggctc	tggggccttg	tgggtttatg	gattcccagt	gctgagtc	9360
ccctcacagg	ctcttggtgg	caccttggac	attggtcaga	agcatgtggt	ccccgggaac	9420
acaccttttc	ctgatcatct	gggaagggca	gcttggtgca	gagaggccac	ctgttcagcg	9480
ccacggcccc	ccagacagct	gcagccacag	ccttgccctt	gatcagagca	aacaccagac	9540
atgtgtgtca	tgcccccaac	ccatctccag	gggacacatg	tcctttcttg	ccaggcctga	9600
gatgaacaag	agagggacaa	gtccccaaagc	ctctctctcc	ttcctgcctc	accactccg	9660
ctgttagatt	ctcaagggtg	atgggtgggt	aactagggca	accgaccatc	ctggtttacc	9720
tagaactgag	ggggcatctt	caggaataaa	actgcaaaag	tctggagcaa	acaggagcaa	9780
gttggtcact	ctggggctgg	tggagtcagg	tttcttctg	caggccccct	ccccgcaagc	9840
atgggtggaa	cccaggacag	gaacacagag	caggccccag	gaccgggctt	gtcacttaca	9900
agtccttttt	tttttttttt	ttttgagatg	gagtccttgc	ctgtcatcag	ggctggagta	9960
cagtgggtgcc	atcttagctc	actgcaacct	ctgccttctg	ggttcaagt	atccccctgc	10020
ctcagccctc	tgagttagct	ggactacagg	tggcaccacc	acgcccagct	aattttttgt	10080
atttctagta	gagatgagat	ggccaggctg	gtcttgaact	cctgacctca	agtgatctgc	10140
ccgccttggc	ctcccaaagt	gctgggatta	cagggtgtgag	ccactgtgcc	tggccccact	10200
cacaagtctt	aaacctatgc	tcagcacatc	aatgccattt	acaaaaaggt	agagggattt	10260
tccaggcaaa	aatatagtaa	agacatagga	tgattgatca	tgtcctgctt	aaacataggt	10320
ctgatgctat	taagaattga	gggctgggag	cgggtggctca	cgcctgtaat	cccagcactt	10380
tgggaggccg	aggcgggcgg	atcacgaggt	caggagatcg	agaccatcct	ggctaacacg	10440
gtgaaacccc	atctctacta	aaaatacaaa	aatggccgc	gcgcgggtgac	tcacgcctgt	10500
aatcccagca	cttgggagg	ccaaggcggg	cggatcacga	ggtcaggaga	tcagaccat	10560
cctggctaac	acagtgaagc	cccgtctcta	ctaaaaata	caaaaaaat	tagccaggca	10620
tgggtggcgg	cgctgtagt	cccagcaact	tgggaggtg	aggcaggaga	agaatggtgt	10680
gaacctggga	ggtggagctt	ccagttagcc	gagatcacac	cactgcactc	cagcctgggc	10740
gacagagtga	aactccatct	caaaaaaaaa	ataaataaat	aaataagaat	tgtagtatt	10800
ttgcaggtgt	gacaaatgat	tctgtttctg	tggcagaatg	ttctcaggag	atctcttttg	10860
aactctcatg	gaaagcatca	tgctgttggc	aacatcacat	ttatttttat	ttatttatta	10920
tttttttagag	acagggtctt	gctctgttgc	ccaggctgga	gtgcagtggc	acaatcacag	10980
ctcactgcag	cctcaacctc	ctgggctcaa	gcaatcctcc	tgctcagcc	tcccaaagta	11040
gctgggacca	caggcgtgag	ccactgcact	cagcccaatg	taccttcaat	atttacattt	11100
ctggcaaaag	tagcaaaacc	ttaacaaatt	ttgaatctag	ataataaaat	tatgaggctg	11160
ggtgcagtgg	ccctgacagg	gatggctcac	atctgtaatc	tcaacatttt	gggaggccaa	11220
ggtaggcgga	tcacctgagg	ccaggagttt	gagaccagcc	tggccaacat	ggtgtaaccc	11280
tgtctctaac	aaaaatacaa	aaaaatttagc	cagacgtggt	ggtgcacgtc	tgcatccca	11340
gctactaggg	aggctgaggc	aggagaattg	cttgaacccg	agaggcagag	gttgatga	11400
gccgagatcg	cgtcattgca	ctccagcctg	ggcaaaagca	agagcgaaac	tctctctcca	11460
aaaaataaaa	aaaaataaaa	ttaatgaatt	aattaaaata	aaataaaaata	atggatagtc	11520
actgtaaaga	aaaaataaat	gtatatatca	gccaaacaagt	gatggaatag	agcaccat	11580
ctccctggct	ggacagatac	atcccacaac	acctggaagg	cggctccatg	tagaactttc	11640
tggactgctt	gaggtgctgt	gctggagcac	ggtgacagag	gagctggacc	atggacctcc	11700
cccggccccc	accaagggag	aggtccccc	gtgggtgggtc	tgaggagggc	atccgtatgg	11760

cctctgcggc	ttgggcaggg	aatttggggt	ccaagtactt	ggtgcaaagc	ctggaaagag	11820
ggtttgggtg	ctgagggcat	atcccctggg	ccacatgggg	gcagaagtgg	ggccccctga	11880
agcttgaggt	cctgggcagg	ggcatctatt	ttgctgtctg	aggccttcag	tacttgaagc	11940
aaaatggagg	cagaatgtcc	caccttaatg	cccctgattc	ctccaaacca	attccagaga	12000
cagcaagggc	cagaacaggg	atggccctgc	ccagggtcac	gcancgagga	agtggccagg	12060
ctgggatctg	aaccagggt	aatcccctcc	cttgctctcc	tccaggccct	cacccttgca	12120
tagagccctc	cagctcactc	atcctcggcc	agctccatct	cctcagcttg	taaaccccc	12180
cgggattttc	ctttcttaaa	aaacaaaggc	ttggccaggc	acggtggctc	acgcctgtac	12240
tttgggggtg	gctcccagca	ctttgggagg	ccaaggtggg	cggatcatga	ggtcaagaga	12300
ttgagaccat	tctggccagc	atggtgaaac	cctgtattta	ctaaaaaaa	aaaaattaac	12360
tgggcatggt	ggctagctac	ttaggaggct	gaggcaggag	aatcgcttga	acctgggaga	12420
aagaggttgc	agtgagccaa	gatcgcgcca	ctccacttta	acctggcaac	agaacaagat	12480
tccgtttcna	aaaacaaaca	aacaaacaaa	taaacaaaaa	aaggcggagc	gcgatggctc	12540
gcgcctgcaa	tcccagcact	ttgggaggct	gaggcgggag	gatcacttga	ggttaggagt	12600
ttgagaccag	cttggccaac	atggtgaaac	cccatttcca	ctaaaagtac	aaaaatcagc	12660
caggtgtggt	ggtgggtgcc	tgtaatccca	gctactcagg	aggctgaggc	aggagaatcg	12720
cttgaaccca	tgacctggag	gctacagtga	gctgagattg	cgccactgta	ctccagcttg	12780
ggcaacaaga	tttgtttctc	taaaaaaaaa	aaaaaaaaaga	ctggcccttc	cccttcagct	12840
cttctcagg	gtccctgagc	actctacacc	cccgtctaca	ctgagcactc	caccctgctg	12900
tctacactga	gcactccacc	ctgccatcta	cactgaggac	tccacccccc	tgtctacact	12960
ggctgcctcc	cgccctcacc	tcctgctaag	gccattcccc	gctgcatctg	tcttctagat	13020
tctgcagcct	tcagcacgct	gggccccctc	tttgtccctc	tgagccacct	ccagcctccc	13080
cctgagctgc	tactcctctc	ccagcagcct	ccacccaagc	ccctccagtc	cccaagctgt	13140
cccttgcatc	cagcactgcc	cttccacgtg	ccccttccct	ccagcttcac	agcaggggtg	13200
ggcctccagg	ccctgcccac	tgtgcccatc	cacaagttgt	ggtgggagct	ccgaggggag	13260
gcaggggtgt	gcatggactt	gggacgtcca	agtctgggac	caggggagc	tgggttggtg	13320
agtgtggagg	gggataggga	ctttcaggta	gagaggctgt	aggggcaaga	tcgggacggc	13380
ggatgtccct	aaggagggtc	ctgacctggg	aaatattgtg	cagcttcctc	tttgccattc	13440
ctggagctca	gacactggcc	ggctctcacc	ccgccttccc	tgcaggacac	agctccatcc	13500
cagtgagttc	ctagtgtaga	catctccagc	agcacggatg	ggaaaggaag	tcatcaaagg	13560
tgcccaggac	cggaggcttt	ttctggaggt	ggcagaggag	ggtgtgggtc	tcagggctct	13620
ggctgagggc	aagcgtggga	ggtcttaggt	ctgcaccagc	cccgtgaagg	cccctcctgc	13680
tccctggtgg	agtccatagag	ggaacagcag	cccctaggct	ctagcaggag	tgggtagggg	13740
cttttctggc	ttcctactgt	gccagcagga	tagctgggac	tggcactgag	cccaaagatc	13800
acatgccggg	gcattggcgc	agtgaggaac	agacccttgc	caaagctggc	aaagaagacc	13860
ccatgggggtg	cagctggtga	agctgagagc	tcaatgtttg	ggggagcctg	gcaaaagggg	13920
tcctccctcc	cctctgcagg	ccaggatcgc	aggttttccc	tacatgttgg	taattctcaa	13980
acaatcccat	ggccactgga	gcaaagatca	cagtggcgcg	cggcctcggg	agcagtggac	14040
agggcacgca	tgtgccttga	tgccagagcc	ctgcgcccaa	agtcaacaaa	ctctgcagcg	14100
gactttgcac	ccggactttg	ttttcaccat	acaaggaaaag	ggacagatca	caggccctct	14160
cgctgccctc	gctgagccgg	aagctgcagc	gtgagctctc	tcaagcccca	tttctaggtt	14220
ccccaggcgc	acccctgagc	ccctactcgc	ctattaagtt	ctcctaatag	cccttcaagg	14280
tcttaatgta	tgtccattag	acagagggga	aaactgaggc	gagggcaagt	gacttgaccg	14340
aggttcctcg	gcgagcaggg	cgtggagctg	agaacctcgt	tattactgct	ccccacacaa	14400
ccctctggcc	gttcttggaa	gaaggctgag	ccccgggggg	gccagagtga	cccaaacc	14460
atgggcccgc	tgcggtaca	cgtgcggcca	cgaaggggca	gcagtttccc	gcccggccgg	14520
gctctctccg	gcgctcagta	tccgtcccag	gccaaagaaga	agaaactcgg	ggaggagggc	14580
ggaggggggt	gcgtgggagg	gcgtggaaga	tggacgtggc	caggggagtg	gcagctgcac	14640
acagtggatg	ctgttaagat	gaagggaag	aacgtgggct	ccgagatcac	tggacacggg	14700
tccacctttc	ttcccgtcca	ctgcatggcc	ctgggcgggt	tgttgaaccc	ttggaaaacct	14760
gtttttcctt	ttttcctttt	tttttgagac	agggtcttgc	tctgtggccc	agactggagt	14820
gccgtggcac	gatcttggct	cactgctgcc	tcccaggttc	aagtgatcct	cccagctcag	14880
cctcctgcgt	agctgggacc	ccaggtatgt	gtcaccacag	ccggctaatt	tttgtatttt	14940
tttgtagaga	cgggatttcg	ccgtattgcc	caggctgggtc	tcaaactcct	gagttcaccg	15000

gatcttctctg	cctcagcctc	ccaaagtgtc	gggattactg	gcatgagcca	ccgcacccag	15060
cagagacctc	agttttctaa	cctgtgccag	caggaataat	gatagctgcc	tagcttggct	15120
gtgctgggaa	ttaagtaaga	tgaccgggta	gcaaataatga	agtattactg	gacacagagg	15180
gccccaggct	gggttagcag	cggtggtcag	ggctgctgct	tcctggcctg	agctcgaagg	15240
agggccctca	ttaccacctg	ggtgagtcct	cgccaagcc	tggcactgct	gcgtgggaat	15300
aacttctgcc	acccaagttg	gcagattgtg	tgcaaagtta	agtctgact	ctgtgggggtg	15360
gacttcagag	cctcttcac	ggacctgctt	ccggtgactg	cattcgcacc	tcctcctgtt	15420
cctggtttaa	cacagcccag	ctttcctcct	gctgagccct	ccctgggcct	gctgtcaccc	15480
tcgtgccgct	gtgcctcgca	gtgccactcc	ctgtaccctg	aatactttgc	cctgcctctc	15540
caccagctg	agagtcaggg	cccctgtgag	gctctgcccc	gcccgtcctc	cgggtttctg	15600
cctctgctga	gcacttccct	gcatgattgc	ttctgagagt	ccccccagcc	tgtgagcttc	15660
tcaggactgg	gacagcttct	caggaccgag	gcttcctggg	ctgcttgcaa	ttttacaggc	15720
gggcacattt	tcccttggcc	aacatcagag	actggacatc	tgcatatctg	tgctagccac	15780
tgagcaccca	ggcaccccag	caggtagctc	tgtaaccaac	ccattctgta	aagctgaggc	15840
tcagagaggt	gaagcgccctg	gcctggggcc	acagcctgcg	tcagctgcag	agccaggagc	15900
tgagatatgc	acctgcggct	ctgctcacag	ggctcctgcac	agactgctgc	tggagccacc	15960
tatgtagagt	caagagaggt	catgttaact	ccctctcaca	tcctcagcc	aggggtggggg	16020
ctgacgatag	acactcaggg	atggcctacc	ctccccaaca	accccgtca	ggtttgccgg	16080
atctccttgg	aagaaaagtt	ctgggcagaa	ttccaccggt	ggcctggcct	acactctcct	16140
tagtggttta	ggaccctcag	cgggtggataa	gttgtgggca	gaagagatgc	aatcaggatt	16200
ctcaccact	caccocctgc	cagccccaat	aagctcaata	agctgggctc	ggtctgagga	16260
agtgtccagg	aaatgtgcaa	atggcctggg	acagccctgt	gttcctttca	gtaagggtgc	16320
tgaaggtgag	gctgaaagtt	ggagaaacag	aagccagtgc	ttatggtttt	aattaagata	16380
atggaatgta	tgtatgtatg	tatgtatgta	tgtatgtatt	tatgtattta	tccttagaga	16440
tagagtctca	ctctgttgcc	caggctggaa	tgcggtgaca	caatcatagc	tccttgacgc	16500
ctcgacttcc	tatgccccaa	tgatcctcct	acctcagcct	cctgagtagc	tgggactaca	16560
gacacacgcc	aactatgcct	agctaatttt	tatttctatt	ttttgtggag	actgggttct	16620
cactttgttg	cccaggctgg	tcttgaaacc	ctagcttcaa	gcaatcctcc	tgccctagcc	16680
tcccaaagtg	gagggtattac	aggtgtgagc	caccacacct	ggcctggaat	ttatttgtat	16740
tctgcttata	aaattaatac	attcttattg	cagaaaagtt	tgaaaataaa	agaaaggaca	16800
aagaacaaaa	agcgtatata	atttcacagc	tcagatctca	ctgctattaa	catttttatt	16860
tactttcagg	cttttttctt	tctaggtaca	tatgcagaga	ttattttatt	ttattttatt	16920
tattttatat	tttattttat	attttttatt	tcattatttt	atttttattt	atttttattat	16980
tttttagagac	agggcctcac	tctgtcaccc	aggctggagt	acaatggagt	gatcatagct	17040
cactgcagcc	tcaaacacct	gggctcaagc	aatcccccca	ctcagccttc	tgagtagttg	17100
ggactaaagt	gtgagtctgg	ctaatttttt	ttactttttg	tattgacaga	ggtctcacta	17160
tgttgcccag	gctgatctca	paactcctggg	ttcaagcgat	cctcccacct	tggactccca	17220
aagtgtctgg	attacaggca	tgagccacca	tgccctggcct	aaaatgccac	tttttgtcat	17280
ttactaaaaat	cccatggaca	ctttgacatg	tctgtattct	atgctattga	tctgactggt	17340
ggcatctaca	tcattatggc	catctatcat	ctatcataat	ccattttaac	attaaaattg	17400
tgctgtctgct	tagatttttc	tggcctgtct	cctattttgta	ttcttccaga	taaaatttag	17460
aatcatttta	tcaaatcccc	cttgccagaaa	aagccctatt	ggattttggg	tgaaaaatac	17520
tgaatttttta	cattaactta	ggaaagggct	gggcacgggtg	gctcacgcct	gtaatcccta	17580
cacttttcga	ggccaaggca	ggtggatcac	ttgagggttg	gagtttgaga	ccagcctggc	17640
caacatgggtg	aaactcggtc	tttactaaaa	atacaaaaaat	tgccaggcgc	attggctcac	17700
ctgtaatccc	agcacttttg	gaggccgagg	tgggtggatc	acgaggtcag	gagatagaga	17760
ccatcctggc	taacacgggtg	caaccccgtc	tctcctaaaa	atacaaaaaa	ttagccaggc	17820
gtggtgggtg	gcgcctgtgg	tctcagctac	ttaggaggct	gaggcaggag	aatgggtgta	17880
accaggagg	cggagcttgc	agtgagccaa	gatcgcgcca	ctgcactcca	gcctgggcga	17940
cagagtgaga	ctccatctca	aaaaaaaaata	ataataataa	tacaaaaatt	agccgggggt	18000
cgtggcgctg	acctataatc	ccagttactt	gggaggctga	ggcaggagaa	tcgcttgaat	18060
ccaggagggtg	gaggttgcaa	tgagcagaga	tcgtgccact	gtactccagc	ctgggtgaca	18120
gagtgcact	ctgtgaaaaa	aaaaaaaaaa	ttctgaaggga	ttgagactct	tagactctta	18180
ggtcttccta	tccaagagca	caatatagct	tttcatgtat	tcaagccttt	ttcaatgcac	18240

caacagaatt	ttacagtttt	tttcatgata	tcctgctatt	tcttataaaa	tgtattccta	18300
gataattctgc	atgttttccg	gttgtttgtt	aataaatatt	tttcatttgt	cattatttcc	18360
taattggctg	ttatttgtat	atatgacatc	tgttgaattt	tttgattact	ttgaaaatgg	18420
ccattctttt	gtgttttttt	ttaactttct	attttgagat	aattttgact	tacagaagat	18480
ttgcaaaaat	agtacagaga	gttcctgttt	cccccttatg	ttaaccaggt	ttctccttat	18540
gttaacatct	tacataacta	cagaacaatt	gtcaaatcta	agaatcaadc	tgggcacaat	18600
gctattaact	aaactgcaga	agctgttcag	atctcaccag	ttcttctact	gctccccttt	18660
tctcttccag	tgttcaatcc	ggaatcctac	attatattta	gttgtcattt	ctctttgggtg	18720
tcttccaatc	tgtgacagtt	cctcagtcct	tctttgtctt	tcatgacttt	catttttttta	18780
tactttttgaa	aaatactggc	cggttgtttt	gtagaacgcc	ctcagtttgg	gtttgcctga	18840
agttttttgt	gattagatcg	aggtcatgca	ttattggaga	gggtgccacc	gcctcgatgt	18900
gcaagctcaa	tgcacatcat	cagagggttt	gtaatgtcag	tttataccgc	cggagaccct	18960
aacctggagc	atttcgtgaa	ggtgctgtct	gccaggattc	tccactagaa	agttactatt	19020
tttccccttt	taattactga	atgtctgagg	ggaaatactt	tgagactatg	caaatatcct	19080
gtttctgctt	taacttcggc	tcactaagtt	tagcattcat	ctatggatct	cgcttatagc	19140
aagtattact	gtggagtctt	aatggtaatt	ttctgtttct	ctcattcctt	caacccttat	19200
taatatgctt	cttcctcact	tattcatttt	gtttcagttg	tttataccaa	catggatttg	19260
tggatattgg	ttttattctt	tgggttgcaa	ttgaaacctt	tcattatttt	gttagtcagt	19320
tgttccatcg	gacctgtggtc	attaggagcc	cttgaaattt	ggctcccatg	cctttttttt	19380
tttttttgag	accgagcttc	actctgtcac	ccaggtttga	gtgcagtgcc	atgatcttgg	19440
cttcctgcaa	cctccgcctc	ccagggttcaa	gcaattctcc	tgctcagcc	tcctgagtag	19500
ctggatttat	aggcgctcca	ccaccttgcc	cggctaattt	tttgtatttt	tagtagagat	19560
gggtttttat	tatgttggcc	aggctggtct	caaactcctg	acctcaggtg	atctgcccgc	19620
ctcgccctcc	caaagtgtctg	ggactacagg	cgtgagccac	cacacctggc	ctcctatgcc	19680
attttaacat	gcccgtcttt	tctttttctt	tcctactttc	tgtgactgta	agaagctcca	19740
ggatacattt	ttgctgccct	agacttagcc	tcaatcagtt	ctcagaaaag	ctctgggttct	19800
ttttatggga	tacttagaaa	actagctctg	tatggcctgg	cgcggtggct	cacgcctgta	19860
atcccagtac	tttgggaggc	cgaggtgggc	agatcacaga	tcacgaagtc	aggagatcaa	19920
gaccatcctg	gctaacatgg	tgaactctg	tctctactaa	acatacaaaa	aattagtcca	19980
ggcgcggtgg	cgggcgctcg	tagtcccagc	tactcaggag	gctgaggcag	gagaacggca	20040
tgaaccggg	aggcgaggct	tgcagtgagc	cgagatcggc	agccactgca	ctccagcctg	20100
ggccacagag	cgagactccg	tctcaaaaaa	aaaaaaaggga	aaaagaaaaa	agaaaactag	20160
ctctgtatgc	tagttttttt	tttaagacag	ggtctctctt	gccccagctg	gagtgtagca	20220
gcacgatcac	agctcactgt	agcctcaacc	ttctgggctc	aagcaatcct	cctgcctcag	20280
tctcctaagt	agctgggtct	acaggcatgc	accaccgtac	gtggcaattt	ttaaaaactg	20340
ttttagagaa	tggagtctcc	ctatgttgcc	tggctggaa	ctcctggcct	caagtgatcc	20400
tcctgcctcg	gcctcccaaa	gtgctgagat	tacaggcatg	agccactgta	cctggcctgg	20460
ccaaggctcg	tcttttttta	aaagaagttg	ttgtatagtt	gttttttttt	ttattttttt	20520
ttctgagacg	gagtctcgct	ctgtcgccca	ggctggagtg	cagtgggtgcg	atctcggctc	20580
actgcaagct	ccgcctccca	ggttcacgcc	attctcctgc	ctcagcctcc	cgagttagctg	20640
ggcctacagg	cgcccgcctac	cacgcccgcc	taattttttg	cattttttagt	agagacgggg	20700
tttcaccgtg	ttagccagga	tggctctgat	ctcctgacct	cgtgatccgc	ccgcctcggc	20760
ctcccaaaagt	gctgggatta	caggcgtgag	ccaccgcgcc	cggcctgttg	tatagttttt	20820
atctcgagtt	ttctagcgat	ttaatcatat	tggttacaaa	aaaggatgat	tttactacct	20880
cctttccaat	gtttctacat	attttttcat	tttatctaac	tgcattttta	aataaaacttt	20940
taatttttaga	atggtttcat	atttacagaa	aatgtgcaaa	gatagtacag	agagttcctg	21000
tgtactccac	acccgggtttc	cttattatta	tcttaacgtg	atacacaatt	aataaaccag	21060
taacattatt	attcactgaa	gtccacactt	tctttttttt	tttttctgag	acggagtcta	21120
cttctgtcac	ccaggctgga	gtgcagtgcc	gcaatctcgg	ctcactgcaa	cctccacctc	21180
ctgggttcag	gcaattctgt	ggctcagcat	cccaagtagc	tgggaataca	ggtgcccgcc	21240
accacgcccg	gctaattttt	tgtattttta	gtagagatgg	ggtttcacca	tgttagccag	21300
gatgggtctt	aactcctgac	ctcgtgatct	gcctgcctca	gcctcccaaa	gtgctgggat	21360
tacaggcggtg	agccaccgcg	ccggcgctcc	atactttctt	tagatatcct	tccttttttac	21420
ctaacgtcct	tcttctgggt	caggatccca	tccagaaagc	aacattaccc	ctcgccatca	21480

cgtcttcaca	ggctcccctt	gacgggaaga	gttcctcaga	ctttccttgt	ttttgttgac	21540
cttgacagtt	ttgaggagga	ctggatatctt	agtctgtttt	gtgctgctat	cacagactag	21600
ctgagaccga	tacatgatac	atgaaaaaaaa	atgtattctt	acagttgtgg	aggctgggaa	21660
gttcaagacg	aagttgctgg	ttggtttgggt	ctctggtttc	aagatggcgc	cttgctgctg	21720
catcctctgg	agaagaagaa	tgcggtgtcc	tctcactgca	gaagatggaa	gcgctaaaag	21780
gaatgaactc	cctttgccaa	gccattttat	aatgggcatt	aatccacaaa	ggatgaaacc	21840
ctgagaaaca	tcaagcttta	aagcactggg	tctcaacctt	tttggctctca	ggagcccttt	21900
atactcttaa	aacgttttga	ggatcccaaa	aaaaggcttc	tacaggttcc	atcttttaaat	21960
atttaccata	tcaaaaatta	aactgaaaaa	attttaaatt	atttattcat	ttaaaataac	22020
aaggataaac	ccattacatg	ctaacataaa	tcatgtattt	tatgaaaaat	agctatatatt	22080
atcaaaaaca	aaattagtga	gaagagtggc	atgtataatt	ttttttgttt	attttttgtt	22140
tttagatgga	atcttattct	gtcgcccagg	ctggagtgc	gtggtgtgat	ctcggtcac	22200
tgcaagctct	gcctcccagg	ttcacaccat	tctcctgcct	cagcctcctg	agtagctggg	22260
actgcaggtg	cctgccacca	cgcccggcta	attttttgt	tttttagtag	agatggagtt	22320
tcaccgtgtt	agccaggatg	gtcttgatct	cctgaccttg	tgatccaccc	gcctcagcct	22380
cccaaagtgc	tgggattaca	ggcttgagcc	actgctctg	gcctaaattt	ttgtgaatgt	22440
ctttaatgcc	tgcttctca	tatttgtttc	tgcatccaag	ttattgcaaa	atgttgtgtt	22500
ggttgaagtt	tgtaaagaaa	atgtggcctc	atacagttgt	gtagttggaa	aggcaagagt	22560
attttgattc	tctcttcaaa	caactatgga	caacctgctg	ttacaaaacc	agaatgcaaa	22620
aagttgtagt	aaatacaggt	taggtgtagt	gtggaatctg	aaagcatgtg	aatgaacttt	22680
ctgagttttg	taacattaaa	gtccagttgc	gttaagctac	tgtgatagca	tatagcattg	22740
tcctaatact	ggaattagta	tcagaagtgg	gggtgctactg	ttaataaata	aaaagaaata	22800
aataaatcat	gtgatactgg	ctcagaagtc	aggcagtagg	ctgtgtggaa	cctgacatca	22860
cgccatgtaa	tacattggca	accatttgat	ccagctgtct	gtcatgatga	cttggaaggt	22920
caaccacata	cttacagagc	ctgtagacat	aggggaaaat	agtataaaac	agaatactaa	22980
cagtggacct	tggttcttgc	cagttgcatt	tagccaaata	ttaaacaaaa	gagatattct	23040
tgggcagcaa	ctggaccatc	ttcaagtaaa	agtgaagggt	aataaacaga	gtccagacat	23100
ttgtgcccct	gcgggttaag	aaaaatccag	ttgttcttag	acaccgtata	tgaaaacaac	23160
gctgaaaaca	agccttttag	tggtaaaggc	cgattaacac	tcagcgcggt	aacaaagacc	23220
aggtgggcta	acccgaaatg	aaatgagaag	cctgtgggtga	tgaggaggca	gagaagtaaa	23280
atcaagtttg	agcatttcgt	ttaggagagt	ttgggctctg	attacttgca	catgcaaacg	23340
aactggaaac	aaacagatca	gatgtctacc	acttcttcga	gggaattgca	ttgccaaaga	23400
agtcatgaaa	gcagactcta	tactgattag	gcattaaaac	aaaaacaatc	tttaggcccc	23460
taaacttgca	tgggcaggaa	gtgggctgtc	aaagctgttc	atcctctaag	gtggacctag	23520
ttcctagtcc	ccagtataca	cttcagatgt	ggccctggag	gacactggac	atggaggacc	23580
tcccagagga	tgaggctagg	gcttcatttc	tccaatgacc	tcagctgcct	ctatttcccc	23640
ttcttccctc	ggaagtccca	tcategttat	tattattatt	atcatcattt	ttatttttag	23700
ataaggtctc	gctctgttgc	ccaggctgga	gtgcagtgc	atgatcatgg	ctcactgcag	23760
ccttcccagg	ctcaagtgat	cctcctgcct	cagcctcctg	agtagctggg	agtacaggca	23820
catgccacca	tgcttggcta	tttttttttt	cagtagagat	agggtctca	ctatgttgcc	23880
agggtctgat	tcaacctcct	gggttcaaga	gatcctccta	cctcagctcc	tgagtagctg	23940
ggattcgggt	gcacaccacc	atgccaacta	attttttaatt	tttttttcta	tggacaggat	24000
gtacagtgtt	agaaatggat	tgcttgca	ggcaggagga	tcacttgagc	ccaggagttt	24060
gatcacactg	tgaaccatga	tcgcaccctc	gcactccaat	ctgggcaaca	gagtggagacc	24120
ttgtctcaaa	aaaaaaaaaa	aagagagaga	gagagagact	caaagatagg	caaaaaagtg	24180
ggaaagcttt	atagtggaca	aaaaggaacg	ctctaagtc	gccctatttg	catggtgctg	24240
aaggtgggct	aactagagat	aggggggtact	atgtgggtga	ctatgggtgc	atctttggct	24300
ttccctgggt	gatcctaagt	tggaagcagg	gacaaaaatt	agggaagctg	ttagttattc	24360
atcacgttct	ggcagtagtg	gactgggtgt	gatagaagtt	attgttttgg	ccagggtcgg	24420
tggctcatgc	ctgtaatcct	agccctttca	gagttcaacg	tgggtggatc	aggaaggagg	24480
gaggatttgg	gaggtcagga	gttagcctgg	ctaacctggc	gaaatcccat	ctctactaaa	24540
aatacaaaaa	ttagctgggc	gtggtgggtg	atgcctataa	tcccagctac	tcgggacgct	24600
gaggcaggag	aatcagttga	acctggggag	gcggagggtg	cagtggagcca	agatcgtgcc	24660
caatttcatc	tcaaaaaaaa	aaaaaaagtt	atcgtttagc	ttcctcgatt	gttactggac	24720

gtagtaatct	ggcttcctgc	aagtctaact	ttcagcagac	tggtacatg	ggctgtgtac	24780
tgtagataag	gcagtaagta	aagcaaaaat	tgatagagca	tcaaggataa	atagaaaatc	24840
cgtaatcaag	cagaagatctt	gaacacttca	ctttcagtaa	ctgataaaac	aagtagacaa	24900
aaaaaatcag	taaggatgta	gaagatttga	acaacgtaat	taacaaaactt	gacttgattt	24960
acacgtctag	aacctgcag	aacacacact	ttttcaagca	tactcagaac	atztatataa	25020
agtgaccata	tggtggacca	taaagcagtt	tcaacaaatc	tcacaggagt	aaaataacag	25080
accgtgtttt	ctgaccgtaa	gtacagttaa	cctagaaatt	gaaaacaaaa	agctagaaaa	25140
accccatgta	tctggaaatt	ttaatatata	ctttgaaata	acaaatggat	cagagattaa	25200
ttcaaatagg	aatttagaaa	taccttgaac	tgaaaaataa	tgagaatact	ataccccaaa	25260
actgtggggg	gcagctgaac	agtatataga	cgaaaagtat	actcatatgt	gcatacctta	25320
aggagcgggg	aggattgaaa	gttaatggga	ggcaaaagca	ggtggatcac	ttgagggttag	25380
gagttcaaga	tcagcctggc	taacaggggtg	aaaccccatc	tctactaaaa	atacaaaaaa	25440
ttatccaggc	gtagtgaggc	tgaggcaaga	gaatcggttg	aaccaggag	gcagaggttg	25500
cagtgagccg	cgattgcgcc	actgcacccc	agcctgggag	acagagcgag	actccatctc	25560
aagaaagaaa	aaaaaaaaag	aaaaggccag	gcgcgggtgg	tcattgcctgt	aatcccagca	25620
ttttgggagg	ccgaggtggg	cggatcacga	ggtcaggaga	tcgagactat	cctggctagc	25680
acggtgaaac	ccgcctctta	ctaaaaatac	aaaaaaatta	gccaggcggtg	gtggcgggtg	25740
cctgtagtcc	cagctactca	ggaggctgag	gcaggagaat	gtcatgaacc	caggaggcag	25800
agcttgcaat	gagccgagat	cgcgccactg	tactccagcc	tgggcaacag	agagagactc	25860
tgtctcaaaa	aaaaaaaaaa	gttaatggga	ttaacatcca	tctcaagaag	ttagaaagga	25920
atgacaaata	aaccaaaaaa	aaaaaaatca	aaagaagaaa	atcataaggt	caagactata	25980
aagagagtgg	ctgggtgcag	tggtcaggcc	ctgtaatctc	agcatttttg	gaagcagagg	26040
tgggcagatc	acttgagccc	aggagttcaa	gaccagcctg	agtaacatag	agagacctca	26100
tctttgctga	aaataaaaaa	aaaaaattag	ccaggcatgg	tggtactgag	gtgggaggat	26160
cacttgagcc	taggaggttg	aggctgcagt	aagccatgat	tgtgccactg	cacttcagcc	26220
tggttgacag	agtgggaccc	tgtctctaaa	aaactaaaaa	aaggctgggc	gcggtggctc	26280
aaatctgtaa	tcccaccact	ttgggaggcc	aaggctgagg	tcagcagttt	gagaacagct	26340
tgccaacaa	gatgaaacct	catctctact	aaaaatacaa	aaaattagtt	gggtgtgggtg	26400
gcattgtcct	gtaatccag	ctacttagga	ggnnnnctnt	ngattatatt	ttctccttcc	26460
taagtgttta	ttggactgaa	ttcagaatga	tgactctcat	tgagctctt	cctgtctcct	26520
aactacagtg	gcttcogacc	ccactctggt	tttcaactta	ccctctgct	gctcatacga	26580
gtagatactt	ccttccttct	ttctcacttg	ttgctcttcc	tcaaccccc	ccgttggtgt	26640
ccctcctct	ttatcttttt	ctcgcgacac	ctgcgttctc	ttgccctctt	atcatccctt	26700
tctcgaggcg	gtccttttct	ttatccagct	taaataacct	ctcctctgtt	tatttggggg	26760
ttgggttttt	atctctcacc	ctccctctaa	tttcttttct	ctttccgcac	ccatcaagcc	26820
tctcgtgggt	tctcttctct	tactctcggg	tccccccct	ctccccctt	ttttttcttc	26880
acccccccaa	gcgctttgcc	ttttttttct	ttgcccttta	ttcccccc		26928

<210> 7

<211> 29430

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (4336), (4345), (4349), (4392), (4447), (4490)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 7

aggggaaggg	cgggtccgt	agctcacacc	tataatccca	gcactttccg	aggagagagg	60
atcatctcag	gccaggagtt	caagaccagc	ctgggcaaca	cagcaagacc	gcattctctac	120
aaaaacttct	tttaaagctt	aaaaaaaaaa	aaaaaagcaa	agaggacagt	tcaggagaaa	180
agcctgtaga	ggcagcacac	taaggaggag	acgcagccca	ggcaccagga	ggggctggcc	240
atgggcactc	actcctccag	caggcgagtg	cccagcacca	gctggccccc	ccagacaccc	300

aggacacggc	ctgaatggct	ccgtattcac	gtgggtggta	ataaacaagc	aatacacata	360
gccataaagg	acaccttagt	aatgtttacat	cataaacgct	gcagatcagg	gaaatgggtgc	420
aggggtgaagt	gggttggggg	gctgcatgct	acatgagaag	tgggtcgggg	ggctgcatgc	480
tacctgagac	agagcaggcc	ttgctgggaa	agaaggagcc	ggcaggcctg	ggcaaagggtc	540
ctgggggtggg	agcacactgg	agcagagtgt	gggggtagca	tggcggtgct	tggctcctctg	600
ggcgccctcc	caccacgtca	tgtgcccatg	tgcccaaggt	ctctcgtttc	acagccccct	660
gaagctcagg	ggtcacagct	acacagcccc	cagatacctt	ggcctgcccc	aggtcattcc	720
atccagtgat	ggacctgctg	acctctagcc	tgacctctgg	gcagcgtaat	ttgagaagga	780
ggagaaggga	gggcaacaga	cctggggcga	tgagggatgc	acaggggtggc	agacacctga	840
ggctgcacct	tggagcctca	gttctgggtg	tgggtggggg	atggacaggc	tgagggctga	900
agcagctggg	cccggccacc	atcacacccc	aggaccacc	agatcaccat	gaaaaaccga	960
atgtcaactg	gcagcccaga	gtgcagaaca	aacctttcag	aaacacgggtg	gtgactgccg	1020
catcatgaac	ataaaataat	tacgccctct	ccccagggat	cacccttgca	ggagtttgct	1080
ccaagaaaca	ccagaaagaa	ggaaaacgtc	tgagtcacaa	tatttgctga	ggccttattt	1140
gtaatagcaa	aaaaaaaaaa	aaaaaaagaa	caatctccag	cggcaggggt	aactagacta	1200
ttgtctccgt	ggaaaggtag	caccaattaa	ctagtaacaa	aatgactgcg	gtaacaacaa	1260
aacgttcgac	atgtcaacac	caaaaaccac	acaccagca	taaccgtgaa	ccatgatttc	1320
tactagaatg	aatggcagtt	atgagaaaagc	accagcggag	acaaagattg	aaaaagtaaa	1380
gggtggcctca	ttagggagac	aagtctctgg	gtaatatatt	gtaatactgg	taaatatata	1440
gtttttaata	tattttttaa	ttccaaattc	catatatgtt	cctatgaagc	tatttctgca	1500
aatatttttt	tcaggaccgt	acatcacaaa	ggcaaaaagg	ccaggtcagc	tctccagctg	1560
agagtgacca	cttcagagca	gacggcagac	tccagggtta	gcaagcctgg	ctgagacctg	1620
gcccagatac	aactactaac	ccctctgacc	tcaacatcct	gtctgtgaaa	tggggataat	1680
tactgcacct	ccacatcaca	gagtgcgagg	cttaaacagg	atgcttcata	gaaaagcgct	1740
caagaggtaa	cagccgggag	ggggtagtgg	ttttcattaa	ttaaagtgtg	ccttcattcca	1800
gccctgggccc	agctccaaca	caaaacacac	accatccact	cagactcagt	tgcctggatt	1860
caaagcccgg	cctggcctcc	agctgtgaga	ttccgggcag	gatttcccat	ctcccagagc	1920
ctcagtttcc	tcattcatga	aacaggaagt	gatcattcct	tttattttta	tttttatttt	1980
tatttttgaga	cggagtttca	ctctagttgc	ccaggctgga	gtatgatggc	gcaatctcag	2040
ctcactgcaa	cctcggcctc	ccagtttcaa	gcgattctcc	cacctcagtc	tcctgagtag	2100
ctgggattac	aggcacacgc	caccacgccc	agctaatttt	gtatttttag	tagagacggg	2160
gttttgccat	gttggtcagg	ctgggtctga	actcctgacc	tcagggtgatc	cgcccgctt	2220
ggcatcccaa	agtgtcggga	ttacaggtgt	gagccaccaa	gccagttga	caactgcttt	2280
taaagacacc	tctggctgct	gtggaaaaca	gcctggtagt	gcctcaaaaa	gttacacata	2340
gaatgatcct	atgaccagta	attccactcc	tacatatata	cccaaaagaa	ctgaaccctt	2400
ctactcatgt	atgtacacat	acaggtaac	gcagtgttaac	agcagtgttc	acaaagccaa	2460
aacatggaaa	cagctcaaat	gtccataacc	gatgaacgga	taaatgaaac	gtagtctatt	2520
caccacctga	cggaggtgag	agggggcata	aaaaggaatg	atgcataaaa	acgaatatta	2580
tggccaggta	tgggtggtca	cgctgtaat	cccaggactt	tgggaggctg	aggcgggagg	2640
atcacgaggt	aaggagttcg	agaccagcct	ggccaacacg	gtgaaacccc	atctctacta	2700
aaaatacaca	aattagctgg	gcatggtgga	ggcgccctgt	aataccagct	actccggagg	2760
ctgaggcaag	agaatccctt	gaacctggga	aacagaggtt	gcagtgagct	gagattgcac	2820
cactgcactc	cagcctgggc	gacagaccaa	aactccgttt	cggaaaaaaa	agaaaaaatt	2880
agccaggtgt	gggtggcgggt	gggtccctgt	aatccagct	ctacttggga	tactgaggca	2940
ggagaaccac	ttgaaccggg	gaggtggagg	tagcggtag	ctgagattgt	gccactgcgc	3000
tccagcctgt	gtgacagaag	gagactctgt	ctctaaaaaa	caaaaacaaa	aaaggcccga	3060
cgcggtgtct	tacacctgta	atgccaacac	tttggaagc	caaggcaggc	agatcatctg	3120
aggctcaggag	tttgagagca	gcctgggcaa	cacggtgaaa	ccccatctct	actaaaaata	3180
cagaaattag	ccaggtgtgg	tggcacatgc	ctgtaatccc	agctactcgg	gaggctgagg	3240
caggagaatc	gcttgaaccc	aggaagcgga	ggttgcagtg	agccgacatt	gcaccattat	3300
actccagcct	gggtgacaga	gtgagattct	gtctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	3360
ctaaacaaaa	gcaaaaaaac	caatgagtaa	tgttgtcaag	tgaacttcat	cccaatggga	3420
atgcagataa	tttgttttaa	aggcaccatg	cacactgggc	aggctggctt	cccctgggaa	3480
cgtcttcttt	tgcctggatt	cccagttggt	ttaatcgggc	gtagaacact	ttcttcaatc	3540

cgggattcag	gcacccctgc	tcagcacaaa	ctcagtacac	cccgcaactct	gctgtggggtt	3600
cttggcacta	ttaggagaat	gtgagggggt	gattcagatc	tatctctagt	gggtgcatgt	3660
ctgccactcc	caggaacgcc	cacttctggc	aagtcagtgt	cagagaaagg	ccagctcgtg	3720
gccccctctg	ccttgagtcc	caggacccgt	gatcagtcct	acccggagca	gaatcaggag	3780
tttgaaaacc	caagtgccaa	caatctcatt	ttaacccatg	taagcatatc	caatatttat	3840
atatagaatt	cataacagat	gtctgggctt	ccattccaat	agcctatatt	ttacactggt	3900
tatttacatg	gttacaccaa	acaagactca	attcaaggta	acccaatcct	ttgctactat	3960
acaaaaataa	gcaacatttt	cagtccatgc	cttatatata	ttcaccaagc	attacactag	4020
gcctccaact	gctcatcgga	gcaagctgca	gcctggacac	aagctagaga	ttaatcagtc	4080
aggaatgatc	ctgcgtccag	tgccagcatg	atggaagaga	cagagaaaca	gaagacatca	4140
gggctccaga	gtcaaggagc	ctgcagggtta	gttgggcagg	atatacacac	atacacacac	4200
acacgcacac	acaaaaccac	ccaagaagaa	aagggtgggat	gaatgcatgg	acaggtaatg	4260
cctggagcct	gggatgggat	aagctgactg	cagggtggccc	aggcaggcctt	cctggaggaa	4320
gaagacctgg	ctgtangtgg	ggtangcang	ctttctaaat	ggggaaaatc	tggtgttggg	4380
tggagttggc	angtttccga	aaagaagaaa	agctgactat	gggtacacct	ggctgttggg	4440
ggaacangca	ggcttcttgg	aagaagaaaa	tctggctgtg	ggtggatcan	gcaagcttct	4500
tggaagaagt	aaacctgact	atgggtggac	caggcaggct	tcctagagga	agaagaccgg	4560
ctgtgggtga	accaggcagg	cttcctagac	agaggaagat	ctggctgcgg	ttagagtggg	4620
caggcttcta	agaagaggaa	gggctgactg	tgggtagacc	tggtgttggg	tagactgggc	4680
aggcttcctg	gaggaggaa	agctggagca	ttgaaaaaca	aacatgactt	ggtgaatggt	4740
gagcatgccc	agccctgata	cccagaggca	attacgcact	caagttactt	aattctactc	4800
acaatgcctc	acaaacaact	tctctgacac	ctaacacagc	tctgggcacc	ttctagcttc	4860
agctcctcaa	agcagttatt	cacgctacta	ccctgcacac	ctcctcacac	cccaacccca	4920
gggacaggag	ttctgccaga	tgccaaagct	cctgatgcca	aagcctgggt	ctgcttccgg	4980
gctcctcttg	gtctaactgt	ccaccccgca	tggcatgat	gtgcaaaaac	aaggctttgc	5040
aatctgccct	gatgcctggc	ggagcgagtc	cctcccgatt	cgtctccttc	agaaacacct	5100
gggctgccct	ggtcctgtta	tacccccaac	acattctaca	gtcagctccg	caagttccac	5160
aaagatcaac	gctggcggtt	ttatggcatt	ttatttacag	tttttacaat	ataaaaaagg	5220
aaggatgcca	cagctcagcc	agcaggacag	acagagatct	atgatgcttc	tgctgcacca	5280
ttgtttgtgg	tcaagaaagt	ctgttttcaa	tgatttatta	aattgtggtg	ggagatggat	5340
ggtggcagtg	gttaccagca	acatgaatgt	tcttaatgcc	actgaacttc	acacttacaa	5400
atgggttacga	cgataagtgt	tatatgtatt	ttaccacaat	taaaaacagg	taaatgcagg	5460
ccggggcacgg	tggtctacga	ctgtaatctc	agcacttttg	gaggccaagg	caggcagatc	5520
acctgagggtc	aggggttcga	gaccagtctc	gccaaacagg	tgaaactctg	tctctattaa	5580
aaatacaaaa	attagccaga	tgtgggtggg	catgcctgta	atcccagctt	ctcaggaggc	5640
tgaggcagga	aaatagcttg	aaaccgggag	gcagagggtg	ccatgagctg	agattgtacc	5700
attgcactcc	agcctgggtg	acaaaagcaa	aactctgtct	caaaaaata	aaataaaaata	5760
aaaataggta	aatgcaaaca	tatggatatg	taatattatg	ggctattatg	agctacaaaa	5820
aagaatgact	tgggactaca	gttacagccc	tcattcagga	atttgtttta	aatgtgggtt	5880
ggtcgctaag	gcatgtacac	aacattttga	cgttcaata	ttcctagatt	tggaacagtga	5940
gcacccctct	aagctggctc	ttctgtccca	gagggtcccca	ccagtcctcc	agaacttctt	6000
tgctttctta	cacaataaga	tgccccatgc	tgggttgta	cctttccttg	ccccagccct	6060
agaaccagct	tcttcgtgga	caagctctga	ctcctttggg	tggagaatgg	tattcagaaa	6120
cccagacctg	ggctctgggtg	tgctcactgc	tacttggggg	cattgcttct	aggcctctct	6180
gctgatggag	gtaggatata	cacgtacagt	cttccctctt	cccagattcc	gtacttgagc	6240
tcgcctactt	gctaacattt	atttatatcc	cccaaattaa	acctcacagc	acttctgcaa	6300
tcactcactg	acttgacagag	tgtgaaaaaa	ctgagtcacc	atcacacggt	ccaaactgag	6360
gtcaactgag	gccacaacgc	cccatcttct	tgctccggct	gtcgagatgt	aagcaagtgt	6420
ccttctctcg	gtctagctag	tgccatgctt	tcacatcac	tgtgcttttt	gtgggcaatt	6480
ttgctgtata	aaatgtcccc	tgccatatag	ctgctgtgta	gtgctcctag	gtgcatgagg	6540
ctgccccacg	ccttacagag	agaatatgca	tgagaggctt	tattcaggta	tgagttatag	6600
cgtagttggc	catgaattca	atgttaatga	atcaacaata	tacagtaa	aagggtgcttt	6660
ttagagacag	ggtctcactc	tgtcacccag	gcttttagagt	ccagtgggtg	gaccttggct	6720
cactgccggc	tcaacctcct	gggtcgaagt	gatcctccca	cctcagcctc	ccaaactggt	6780

gggattacag	gcgtgagcta	ctgcactcag	cctaaataag	gtgtcttaga	aacacacata	6840
agacaagggt	atgggctgag	tgcggtggct	catgcctgta	atcccaacac	tttgggaggg	6900
caaggtggga	ggttcacttg	aggccagaag	tttgagacta	gcctgggcaa	catggcaaga	6960
cctcatctgt	atatTTTTTT	aaatcagaca	ggtgtgggtg	tgcctgccta	tagtcccagc	7020
tactggagag	gctgaggcag	gaaaatggcc	tgagcccagg	aggtcaaggc	tgacgtgacc	7080
catgattgta	ccactgcatt	ccagcctggg	gtgacacagc	aagacgctgt	cttaaaaaaa	7140
aaaaaaaaaa	aagccagggtc	aggtatcgaa	cagttggcaa	aaacgtttgt	acctgagggt	7200
cacaggaacc	tagcccgatg	tttcccctag	gagcaatggt	tcagtattca	ataattcagg	7260
gttcccagtg	actttatgga	gcataacttt	caagaataac	aagaaccaac	tgtacgtgtg	7320
tatgtatact	cacacttttta	ttttattttta	ttttattttt	tgagacagag	tctcactctg	7380
tcacccaggc	tggagtaaaaa	tggcgtgac	tcgactcact	gcaacctccg	cctcccagggt	7440
tcaagtgatt	ctcagcctcc	caagtagctg	ggattacagg	tgtgccccca	caaccggcta	7500
atttctgtat	ttttagtaga	gacggagttt	cgccacattg	gccacgctgg	tctcaaaactc	7560
ctaacctcaa	gtgatccacc	cacctcagcc	tcccaaagtg	ctggaattac	aggcatgagc	7620
tgccgtgcct	agcctacata	cactttttata	cacacatgca	tctatgacta	tttctctatt	7680
tctgtgcatg	tgtgcgtggc	agtacctaca	gtttcagcta	tgtgtctggg	tactgtctcg	7740
tccaagtttg	taagcacctt	ctccaaagtg	caaagcctgg	cttgtgttac	tatccatatg	7800
tttacttatt	tgctcaatca	atttacttat	tagctccata	accagcttcc	catctgtctc	7860
agtagcctct	gctgtcagtc	acctctgcac	cctaccccac	cttgcttccg	gatgtctggat	7920
gccaatcacc	cccgacacct	ctacatagca	caaccctcga	catgctgctt	ctttatttct	7980
tattttatttg	tttgagatgg	agtcttactc	tgttgcccag	gctggagtg	agtggcacga	8040
tccaggtcca	ctgcaacgtc	cgctcctgg	gttcaagtga	ttctcctgcc	tcagcttctc	8100
aaatagctgg	gattacaggt	gccaccacc	acgcccagct	aattttttgta	tttttagtag	8160
agatgggggt	tcaccatgtt	ggccaggctg	gtctcgaact	cctgacctca	agtgatccac	8220
cttggcctct	caaagtgtctg	ggattacagg	tgtgagccac	cgcgctgggt	ctgcttcttt	8280
aaatgccagg	caccaacatt	tgtgcaatgg	ggtgggagga	aagaacaggg	aggagagcac	8340
actgccggcc	cctgcactga	atccactgat	caatctgggg	gcaactgcca	tctccatctc	8400
ctgtcttctc	atccgtgaac	atctactgca	gtcctctcca	atgtccttct	gtaaagttgt	8460
attatgtttt	gcatacaggc	cttgcatatt	agtctctaga	tataatccat	atactttata	8520
taaaattcaa	accacattta	aaaaaataaa	actagcatga	ctataacgga	gtctgcaaca	8580
ttctcacaga	ctttatgata	aaacatgaaa	cttcaaagat	acttaggggtg	gggcaggggac	8640
aatgttttaag	gctgcctgga	agcctcccca	tccttgagcc	agaaagtcct	atctcccctt	8700
caaggggaaa	tgcttgaaaa	agcactgatc	aggctaaaaat	gacaggggatc	agggagtaaat	8760
caaagtacaa	gtgagctgggt	ctcctccatt	ctgagcacag	caaagttcag	tctctccaag	8820
tccaagaatc	atacacctgt	ttgccaagaa	tgaagttcag	gtgtctacaa	gtggctgaaa	8880
atattcattg	ctggggcatt	aacaacattc	ttggcaaaac	cataccttag	cttctcgtgg	8940
aaatttctta	aggtagaaga	aacaggaaac	accagggtc	gcttttatgt	agacagttcc	9000
atgaagccag	ggaccttccc	cacatccacg	tttcaattac	ctgcacgcag	ctcacagtgt	9060
attcaacatc	tacgcgtctc	tcctactggg	gtggcggtgg	ccactcaaac	cctcatgcag	9120
ctacgatgac	cgcaattttg	gcaacataat	ttcatgtttt	tccttggggt	tttaccgaag	9180
tcagtgcac	aattctgcag	ttgtctaaag	attcaaaatg	agggacttga	cattttacaac	9240
aataataaaa	tcttgggttt	cctttaacca	agcacatgtt	ctgcctttta	gagaaagctc	9300
tgcaaactca	agctggagtg	ggatacttgc	tgacatcttc	aagcacccca	ggaatagctc	9360
tactccccca	tttccacctt	ggctgaacca	tctatatccc	accaattccc	ccaacatccc	9420
tccatccgtc	catccatcca	cccaaggacc	tgctaagcca	ggaggtctct	cccctctacc	9480
ccacagcctg	gcctcagccc	acaagggtc	tctctacatg	aatcccaccg	caccagagta	9540
gaccaagtct	cccgtagact	ccaccctgac	cacctccatg	cctccagcca	ttcccacccc	9600
taaaaaccct	ccctgggtctc	tacaccagc	tgtgaataac	ttggctgaat	gtgacctggc	9660
ctcctggacc	caggtgaagc	ccacgtcctc	cgtaagcccg	ccagctcacc	ctgcctctgc	9720
accttacttg	gagagagccc	gcacttcacc	tcctcagggc	aggcatggct	gatgccaccc	9780
agtggaaatct	ggtgcaaagc	agggcccgggt	gcagagcagg	gctgcctgca	gagcaaggcc	9840
ctgggtgctgg	ggccgagcac	ctccaatgct	ggcgtgggaa	ccatccctcc	cattccagggt	9900
gctgtctcca	tcaagaatga	gcgagctgct	gacatttgca	tgacaataat	gaataaatac	9960
catattttgc	ttcaaatacca	gaatagatgt	ggccagggtt	ggcatatgac	tgttgggaaa	10020

ggacagtttg	cctcttccca	aaccaacttg	gattataaaa	agcttttctt	aacgaccaca	10080
agagcggagg	agctcagggg	cagacaaaag	gaaggctggc	tgcagaaggc	gggagagtgg	10140
ggccttcagg	ggcgggtggg	gagagagaaa	gcctggagct	gcacccccaa	ggtctgtgta	10200
catcaggtgc	tacagaataa	caccacctct	tccagcttgg	ccccacctg	ccctctccca	10260
gcccagtcac	ccagacagca	ccccactccc	cacacacacc	tcacatctgc	ccgcctcaca	10320
ctcaccagct	tcggctctca	atgcaacctg	gaacctgccc	ttggcctctc	agctcagcca	10380
ccccatttcc	tggtggcccc	tggcccccca	tcgaattctc	tctaatocta	atgcacacac	10440
ttgcacactc	aaacacacac	acacacacac	acacacacag	cccagaggaa	aaccataatt	10500
gactgaggtc	caggcaagtt	tcccagagcag	ggaccacatt	tcaaagggtca	gggaagcagg	10560
cgaacaggaa	acatacaggg	ggcacgtttg	ggggtggagc	aggaaataag	aaatcacttg	10620
caaaagataa	aaagaaaatg	aggtagctgg	tttcagacac	ctcggagcac	acagaacagg	10680
acaggcgcct	ccgggtcttc	cctcaacagg	gagatggggc	aggcaggtcc	ctgctgctcc	10740
accgcagagc	tgggggctat	ggccttgaca	ccaaggccct	ggggcaggcg	gggaggcagc	10800
tgttctcctg	cctgtgctcc	cgggcagggc	ctggccccac	aagggaactg	gccgaaggct	10860
ctgcttggct	actccggaaa	gtcctgggag	acaagcaaa	gacttgctag	gtcactccaa	10920
acggcccaga	tgtgacaact	gtgaagaagc	cacaccaaag	caagggtgaca	gaacaattgtt	10980
ggtgacgtca	ggttatcagc	ttacgtctaa	ctccacttac	ccggactcac	ccgtaacctg	11040
ccgtctcttc	ccaaccagta	aaggatgcct	aggtagaggg	gcacaaggcc	tggagcataa	11100
ttaccatttt	aaaggctctg	agaagtccct	cggtagaggaa	gcctagttca	ctttctctcc	11160
cctaggattt	cccaactgcg	cctgatcaca	gaacattttt	tcatttccac	tcaggaaaca	11220
tattttgaaa	aaacactggc	tagaggcaga	agtgaatgg	aaaacacaaa	agtaaaactg	11280
aacaggaggc	actgggcaga	gaacggtcag	aggcgccctg	aatcctggac	cgggtggagat	11340
ccccagcttg	gcctgctccc	ctccctgggc	ccagaccgcc	tccccccatt	tcctggataa	11400
gaaggcta	gcgcatcagg	gtgaagggtc	tgcttgggct	acacccccag	gctcgcccca	11460
caccaatcgc	gctcctgcga	gagccagtga	ctttcttgat	ttggctactg	tggaattgtt	11520
tgcaactaac	caccccagat	acagatacaa	atgacaggat	gatcagatgt	aaaggaccca	11580
caggtctctg	tgatacggct	tcctgcagcc	agcatggcta	gtgccgtgca	gaatgagaat	11640
gaccccaggc	aagtccttgc	ctcccagacc	cagaaccca	tggagcccac	cagggctggt	11700
tcacaagcac	tgtctgggtc	gggcagagat	tccagcaaga	ggaggggaaca	tccatgcacc	11760
ggagccagtt	accagaagca	aatcgccctc	tccaaaaccc	aggctattaa	tggagtccac	11820
tgttgagtgg	agctggggct	tagctatgga	atactgcaca	gcagagatct	tcctgagaga	11880
aagcagtttt	ccctgaaagc	catgtgtcct	ccactaactg	tgttttaatt	gggcgaacgt	11940
ctgtatctca	ttgcagtggc	cgcgcatgtg	ctgacaaggg	gctggggggc	gggtggggag	12000
cagaagctca	ggggcctggg	aggggaaggaa	acaggccacc	agggctcccc	agaaggcatg	12060
tatctctctc	acaaacacac	gcatgcacac	acacgtgcac	acatactctg	caagccctga	12120
gttagcaact	gtggaatgtg	accagctcag	tgatcccagg	acaagctgct	agggaaatag	12180
acatttgatt	gatgtctgca	aatgtgcgtt	ttcactaatt	agaaggttta	gggcagagca	12240
gagaaaaata	tgtatttcag	agtcccagtt	tgacctgcca	gaaaccagcc	cattactaac	12300
attcttattt	tcaacaaaat	atagcattct	gattacatac	catcttggtt	ccacgcctcc	12360
tgccttgcca	agcccccgga	agcggcccaa	ggccatggca	aatagtgaga	gaaacagttc	12420
caggggtggag	actgactcag	gggtgtcagt	cagtggggcg	ctgatggccg	gtgggaggcc	12480
agcagtcatc	accctctcct	tgggacagtt	gagtagctct	cccccagggt	catgtggcca	12540
ctcaggttca	tatgggaggc	gagaggagtg	gcagagtcca	ggagagtggc	tccgaagtca	12600
ctgttccctc	caggcctcag	tgtcttcatc	cattaaatgg	gtaggctgag	gtctgggatg	12660
acaaggaggg	cttgcactta	ctgaaaccca	tgggaggctg	ttcgccgatt	tcttttattg	12720
atggaagaaa	acactcgtat	aattcaagta	ccaattaaaa	ggcaggcact	ggaaccaccg	12780
tctgccaatt	cctagttttg	cctataccaa	atttgagcaa	gttaattgac	ctctcccagc	12840
ctcagtttct	tcgtctgtaa	aatgagggtg	gggatggccc	ccagcccaca	gggcagctgg	12900
aaggattaaa	gaaatcaaac	atctcttaga	gcccacctgg	cacactgtga	tacacaacaa	12960
atgttagcta	tttttgtcta	tgaagtctag	attttatatc	ttgggtgttc	taaagcagga	13020
tacatttatt	taaaaacaag	gattttcatt	aaacacgtac	cccacagaca	gcaaccccat	13080
ggagactgct	cttaattcag	gccagtatcg	aaacgactct	aactacaagc	tttatacagg	13140
tctcttgget	gtccttcaaa	tccaactaag	gtgggtacttc	tgaagcactg	tgcacatgtg	13200
tgtgtgcatg	cacacgtgtg	ggaagggcgg	gctcacggat	ccctcaggta	ccccaccac	13260

gcagtctcaa	gtcacaaagc	gacagagcag	ccgaggaagg	tctgtgcccc	actggaccct	13320
cgtgaagcca	ccaactctac	ctctgcgcgc	tgtcctgcag	actgggctac	cctttgggtg	13380
gggaccagca	ttt gatgcaa	gaaaggcaga	cagaaaagga	aaagggcaag	ttcgactcca	13440
gataacacag	acagtaccaa	gccccagggc	ccataaatgc	cacgcagatg	gaagcattta	13500
ctgcgaggcc	acacagcaaa	cgcacggatc	cagggacgga	ggtgcagact	gcggtgcccc	13560
tgagccatga	ccctgcaaat	taccaccatg	ggaaaggagg	ctgccaaacc	ccccgacagt	13620
cggctgggct	ggcacagact	cgtgggtttc	atcgagggtg	gaggagggtg	gacgtcccag	13680
cccctcccc	atgcccactg	cagagggaag	cggccgtttc	ccctgtgtgg	ttacaaaggt	13740
ctcattgttc	ttcctcacag	ggaggaaact	ggaggaccga	gctcagaacg	catttttagaa	13800
ctggcagaaa	agaacatctg	gggaaggaaa	cacatttcag	aaacaaacat	acctttgtac	13860
cagcttttat	tttctttaag	tggtgaaaaa	ataataataa	taaagacatg	ccaaatttat	13920
catcgctcta	caaaatccct	ttattgagca	aaacgtggca	gctctacttt	caaatgatta	13980
ctgttccttg	aaaattgcag	caacgtggat	gccaaggccc	gaaggccgcc	atcagcagcc	14040
aaacaaaaga	tgccacctcg	ggctccgcga	cactgtacca	tgccagggaa	ctggacagat	14100
ttgggggaatg	ccacggtttg	cctttaaccc	cctgcctcct	ggtctcctga	tgcatctcag	14160
aggctaacct	cttttgagga	actggcattt	ccttagttgta	aatatgcatg	tgggtttggg	14220
agctccctgc	aaagtccagt	gttgacgatc	agctttgatt	tccttggaat	caagtttacg	14280
tgctcagctc	ggaagttaag	aagaatttgg	agaagctgag	cactatggtg	ttgcaggccc	14340
tgggtgaact	cttccaccaa	gcattcattg	tggactgaca	gcgtgcgagg	ggctctgcag	14400
gcagggtgcac	aggacgaaac	acattccgtc	cgggggaaac	ctgcaggaaa	gctccctctt	14460
cttcctaagg	tgccgggcct	agcttcatgg	gtccctaccc	tccacgcctg	tcacactttc	14520
tgagtctcat	gtgggagctg	cctctggttc	ctgacttcac	tcagtccctc	taggaggtgg	14580
aactactgtc	accccatttt	acagatgggg	agactgggca	caaggggacc	aagaaaccaa	14640
tgcaaagtca	cacttggtgg	atcagtgaca	ggggagatca	attcccaggt	tctttctgca	14700
agagttaaat	tgttttcatg	ctgcctaagg	gggggcaact	gaaagaccac	tgcatatctt	14760
tgccaaaagg	gtcaagcaca	ggagccgcag	ccagtgggtc	agatccgcag	aggcgctggg	14820
gtgaccctcc	ccatacctgg	agggatgctt	gtccctcctc	ggccttcact	gggtcccctc	14880
atgaccgtgg	cctcccagga	cctcagcaca	atcccgggtc	tgtgctccag	gacaagccct	14940
ccgtcccca	gactgtgagg	aaatggaacg	aagaggggct	cgctgcagcc	cagcaccac	15000
actgcccctt	ctcaggggca	agaaccgtcc	tggaggactt	ggctttggag	ggggagcctg	15060
ggaggccagt	aagtcaacaa	gcctctactg	ctcatgggtg	ggatcccacc	gcaggccccc	15120
acctgctggg	gcgggcaggg	acgggcggca	cagcttggcc	agggcagata	acccccacct	15180
tggccagggc	gaaggcagga	cacgtgggct	ccagcctggc	cccaccatcc	ctgcacaaca	15240
ctgggcaaag	tccacgtttt	cctcaactgg	gtgttgacat	ctgcaggaca	ggggcatgga	15300
ggtacagagc	gctgaagcca	cacagcaacc	taggagcgag	actccatgcc	tccccgggga	15360
cccctcccca	ccatgaggac	catgaaggct	tcccatgtgc	cgcaaggact	ctgggtgtgga	15420
gacacacgtc	tcctacacag	ccaggcctaa	cgctcttgta	actgggtggg	cccacctggg	15480
ctcacagctg	aggggcagg	agctcaaggc	ttcgcagggc	ctgctctcat	cccagaggcg	15540
atggggagcc	acagcaggct	gcaggagaga	gggtggggcc	cctccacttc	agaggcccca	15600
tctggcccac	agactggaga	gcacatctct	cagcaaccac	ggagcgccaa	ctgcgcacag	15660
ggcctggtcg	tcagagcggg	gcaaaggcac	tgaccgtcac	ggccagggcg	aggggaagacg	15720
ggtgggcagg	gaccttgggc	agagggggaa	gaacctggtg	cccaggctgg	ccctgccttc	15780
agcagtgaag	ctgagtgggg	aggcgctgat	gcagggggcc	agaaagggct	gctgggtcagc	15840
cgggaggagc	ccccacaga	ggaagcagcc	agcccagacg	cagatggcag	gggtcccctca	15900
acaatgtcct	ctgaaaagga	gaggcgggga	ctgctctggg	gacacctaca	aatagatagt	15960
cagccctcag	ccccctgcca	tacttctgac	aaagcagagg	cccccagggg	aggcgcaccc	16020
gaaggtacct	gcacctgtcc	cccagactcc	tagagcccac	ctgaccccat	cccaccaggg	16080
ctccagctac	aaaataaatg	ccgaggccag	ctaggcaagg	acgcacactc	ggtaccgact	16140
gaataggctc	cacgttgtca	tgagcgcaac	ccacaggcca	ccaggccaca	ctatgcagag	16200
ctgagatggt	ttcggccaag	cagcctctca	gctgagctga	acaagtccag	agtccccggg	16260
gggtcgtcac	tatggagtaa	caattgcgat	gcgatggtaa	ccctaacagc	taaccgtcac	16320
tgagccaggc	cctgagctag	gtacttttca	acgctgcctc	tctgcagcct	caggacgagc	16380
ctgtgggagc	ataaagatca	ttccctatca	cggatgggga	aactgagctc	tgaagcagtt	16440
aacgtgcttg	tcccagaccg	cagagctagg	agcaggacac	aacagcaggt	caggcaggaa	16500

cggggtgaggg	gggcctgcat	gggctttctct	ggaggctgcg	catacacgca	acccccagga	16560
ccccgacct	gcacctgcag	ctcgctactg	ccccctcagt	gactccagca	aacctcgggg	16620
taggggaagg	aggctgggaa	tacctcggtg	gtccgaaaca	gcagcttctg	cttggaggcc	16680
actgctgcat	aattggttgct	gcccagcaca	ccccaaagcca	cctgtgccac	ctgtggtgac	16740
cttccagcat	gccttgggtga	ccaagctggc	cttaggtgct	gtgggcagcc	aagaatagaa	16800
cagggccac	ccctcctctt	cacactaaca	caaagcaaga	ggcgggcact	tcgactgagt	16860
gcatccctct	agctcaaggg	cctcacggat	cacaggggtc	agggcaagat	cccaattctg	16920
cattcccgtc	tgcctttcat	cctgctctgc	caacaacagc	cagtgaggct	ggggacatcc	16980
ctgaacctgt	ttctcacctg	aaacacatca	taccattgga	ccccagccct	ccgggagagg	17040
ccctaatacc	tgactgtggt	gagatcagat	cactggttaa	gtacccagaa	gggccttggg	17100
caggggctcc	aggggtgggg	ggtgatgggc	gtggtggtat	cccgctctgg	gctatagtc	17160
accctgatgg	aggaggtctg	tggtcagaac	cgggctgtgc	agggcacagg	agcccagagg	17220
gacccccaga	gctcacctgg	tggtctctga	gcagggctcc	ctcaaccctc	agagaaaagc	17280
acagcaagga	ggccgcccag	agcccagcgc	ctagcaccca	gtggcgtgcc	agacctgcct	17340
ggatcctgga	gatctctcat	cacctccaa	gtcagtcatg	cccaaccag	ggacccacag	17400
cccacggggc	cgtgaagggtg	tgctgagtc	aagaaggcct	tcgacactgg	gaagccaagt	17460
ggcacctcct	ggtgtggagc	aggcggaatc	ccaccagcct	ctgctctgcc	agtgggcaca	17520
gctggacgat	gcagagaagg	ggctgttgct	taataaacgt	catttcctta	agaggataaa	17580
acctttcaaa	acagatggaa	atTTTTTTTT	aattaaaact	ggtggccaaa	gagatggaaa	17640
gcaccccttg	tgcctccctc	ccatcgtgac	ccatcctctg	cacacctcaa	gctgttcgct	17700
gcccaggtgt	ctcctgaggc	actggggggc	ggtgagaatc	cgtgagccct	cggccagccg	17760
tggtctctctg	gagctctgcc	ccaggccatc	agggcacacg	ccgggcaccc	tggggggccac	17820
acagggcaga	gcccagctgg	gtcagcacac	agggccacac	tgggcacaca	agtctctgag	17880
cctccctctgt	ggacgcagct	ctcactatcc	caccccacta	ggtcccgggg	atctgtccca	17940
cagggtgata	tgctgtcaca	gaccactacc	agagccatgg	cctgctgttc	cgcccgagc	18000
caggtagtca	cttgtctccac	agggacaggc	aacgcgcgac	ttgggggctg	ctctgcggca	18060
ggactagagc	tccagcagct	cagccctcct	gagaaggaga	actccatgct	ctaagaggca	18120
gacgcagcgg	acggcaccaa	agccaccaca	agcccacggg	gccctgcatg	gcaggtcagg	18180
agtccctgac	cactcgctct	ttgtaaccag	agctgcagtg	gagtctacga	ggcaaggact	18240
gtgggcggca	gtggccacag	caaataaatg	agtgtcccaa	gggagcaggc	ggctgcgggg	18300
aggcacagcc	gggacccagg	agtcctccgg	cactgcagca	aaactccctgg	gccccctgag	18360
cagcgaccag	gtggcaagtg	catgaactcc	cgggggcata	acctgggagg	gtgacactct	18420
cttcgtgttc	aaattcttga	gaacgcatta	aaaatatcac	tcagtcacct	actctatagt	18480
tttaactcaa	aagtaccaa	gtagccaggc	gcgggtggctc	acgcctataa	tccagtagt	18540
ttgggaagct	gaggcaagag	gatcacttaa	gcccaggagt	tccaaatgaa	cctgggcaac	18600
atggagggac	cccatttcta	caaaaaaagt	gttttaaaaa	attacctggg	cctgggtggtg	18660
tgtgcctgta	gtcccagcta	ctcaggaggc	tgaggcggga	gaaccacatg	aaccagggg	18720
aggtagaggc	tgcagtaggc	tgtgatggca	cactgcact	ccagcctggg	taacagagtc	18780
agactctatc	tcaaaaataa	tttaaaaagc	accaagccag	gcttgggtggc	tcacacctgt	18840
aatcccagca	ctcagggagg	ctgaggcaag	tggatcacct	gagtcagaag	ttcgagacca	18900
gcccagccaa	catggtgaaa	ctccatctcc	actaaaaata	caaaaattac	ccaggcgtgg	18960
tggcgggtgc	ctgtaatccc	agctactcag	gaagctgagg	caggagaact	gcttgaaccc	19020
aggaggcaga	ggttgacagt	agccaagact	gtgctactgc	actcaagcct	gggagacaga	19080
acgagactcc	atctcaaaaa	ataaataaat	caatcaaaac	caccaagact	ttttaatata	19140
aacattttatt	attccataat	tccttttttg	catgattaaa	aatgtttata	taaagtttcc	19200
tgaaaatggt	agaatgccca	agtgaaggct	gcaaataccc	aagccccac	cgtggcatct	19260
cacggagtct	gggccttagg	aggctggtgg	gtaccacgtg	gacccgagac	ttcacagtca	19320
agtccctttg	gggtacactg	ggtttcccac	accccagaaa	tatgggctct	tactgcagga	19380
ccatgggggt	cctcacactt	ggcccagaag	ctgtcacata	gccagacagg	tgttctacaa	19440
cctaggctag	agggagctca	tgctccagca	gaattcgagc	cagaggagggt	aaaagatggg	19500
taagatctgc	tccctggaca	gatgaggcct	tggcctcaga	acagttactg	atcatctacc	19560
agacatcaca	ctagaggcag	aggggcgcag	acgaagacag	ccctgtcct	caaggccctc	19620
ccaggttggg	tggaccatgg	aaggttccag	acagatctgg	caagagaagt	gcccacacca	19680
ggggcagaag	atgggcagggt	ctgctcaggg	cggcacggcc	tgccaggcca	aaaagttcca	19740

acttcagatg	ctggagaatg	ggcacgactg	tctgagaaa	ggaaggatgt	gatgaaaact	19800
acttgaggaa	aaattaatct	ggccagagca	taagataaat	gggcaaaggg	gaggttccag	19860
aaagcaagga	gaccaagtaa	aagctgatgt	cattggctct	gaatctaggc	tttctactgaa	19920
tatgcaccgc	agggcctgta	ggtaaagcct	cagagcccag	ggagtctgag	tggaggagag	19980
ggcaggggac	agagctgggg	cctgtgtcta	cagtgtcag	gaggaatagg	catggacgtc	20040
agctcggagg	ctccagctga	agtgaggagg	cggccagggc	agcacggcca	cgcccggatc	20100
cagactcctt	ttgggaagca	agttcgctct	gggggaaagt	ttggagaaat	ggcctttacc	20160
cgcagaagca	agccccagaa	catatcttgc	tccaaaacta	tctcgtacag	tgaggacgtt	20220
aagcttcagg	tcccctagag	gagacagtct	gctccttcct	ggggcagaac	ccaaggtggc	20280
cagagcctgg	aaggcaccca	gcacccaggc	tgggtgtgtc	cagcccaggc	cacacgctca	20340
gatagctatt	aatgccccgt	tgagcaatct	cctgagagct	ttgccaggca	ggtaccgcct	20400
ccccatctga	actaatacag	gggtacatcc	caaggaagaa	atgaaagggtg	cccacatttt	20460
gctctgggat	taactaggga	ggggagtgat	aattaactca	gtaattatat	ttgccatcgg	20520
gctaattgcta	aaattagtgt	gcattagaat	ttctttcctg	agcagacacc	ggagtgaagt	20580
gggcagcagg	agtggctcgg	gcaagtccgc	acaaagggca	cctccagagc	cttcacacaa	20640
tgtcagcaaa	accacacaa	gtcaaggccg	gctccactgc	accagcaga	tgaattcact	20700
tccacagcct	gagaccgcca	gctcatcgga	ggccatttaa	aatccagccc	tctgacacct	20760
gctggatatac	accatttacc	gtccccagat	caagagatca	aagggtggaa	cctgatagga	20820
cggctctgaa	gttcaccaca	aaagcataaa	cgtgcaagca	gagccaatac	gtcttttgaa	20880
aaggacaatg	aggtgggaat	ttacataaact	gatcttataaa	taigtctctga	tgcttcagag	20940
atggagacag	cagcattccg	gtacacaaaag	acactcacag	gcagtggagc	acagtgaagg	21000
gtctggaatc	aggacccagg	tgtctgtgga	cactacacat	aaaagagcag	cattttacaat	21060
gaatggatag	gatggaccat	cccaccaagg	tgttggacaa	ctccctattc	actggccaga	21120
cccctacctc	ataccatata	caaaaaaaaaa	aaaaaaaaaa	aaaccagac	agaataatgt	21180
ctgaatgtaa	aacataaaac	agtaacagtc	ctggaagaaa	ataatggagg	atataatttat	21240
aatctggaga	tggagtaaca	agggatagga	aaaaagccat	agggaaaaag	tagagttatg	21300
attatatgaa	gcttcttaat	atctttatga	taatgtacca	ccagaaacaa	ggatgaagga	21360
ctagctacag	accagcagtg	aaacctgaaa	caaacagAAC	aaagaattaa	agtccatacc	21420
aaataaagac	ctcccacaaa	tctataagaa	aaagataaac	aggctggcac	cgtggcttat	21480
gtctgtaatc	ccagcacttt	gggaggcgga	gatgggtagg	tcacttgagg	tcaggagttc	21540
gagaccagcc	tggccaacat	ggtgaaaccc	tgtctctacc	aaaaatacaa	aaattagcca	21600
ggcgtgggtg	cgcagtgcctg	tagtcccagc	tacttgggag	gctgagccag	gagaacagct	21660
ggaacccggg	aggcagaggt	tgcaagtgaac	caagatggca	atcgcgccac	tgcaactccag	21720
cctggaggac	acagcgagac	tctgtctcaa	aaaaaaaaaa	aaaagaagaa	gaagaaaaaa	21780
gaaaagaaaa	agacaacaga	aaaatggggc	aaggataagt	gtaggcaatt	tgcaagaaaag	21840
taaataccaa	taaaccagaa	atgagggttg	tgcaaatcaa	aagggtgttat	aatttttaac	21900
caaactggac	caaagaaaaa	accaaaaacc	aaaatcttgt	aattgccagc	atcagagagg	21960
atataggaaa	gtgtgtgttc	tcgtagatgc	ttgcaggat	gaactgctac	agccttttag	22020
gagttatgta	tgtatgtatg	cttgtatgta	tgtatttgag	acagggtctc	gctctgttgc	22080
ccaggctaga	tctgttgagc	tgctgtgatc	atggcttact	gcagccttga	cctcctgagc	22140
tcaatagatt	ttcccacctc	agcctttcaa	gtagctgaga	ctacaggagt	gtgcaatcat	22200
actcagctaa	tttttttaaat	tttttgtaga	catggggggt	ctcccaattt	tgcccaggct	22260
ggtctcgaac	tcctggactc	aagtgatcct	cctgcctcaa	cctcccaaag	tgctgggatt	22320
acctggatga	gccactgtgc	ccggcctcaa	tatctttaaa	aacagaaatg	gacacactct	22380
ttgactagga	atgtatccta	taaaaacact	tatacacatg	cagagacaca	cgagcaagca	22440
tgctttgtaa	tagcaatgaa	ggctggaaaa	actcctcaat	caggtaaatg	ctgtcaagtg	22500
cacctgtgta	ctatgaaatg	gcacttggct	tttaacaaga	gcaaagacag	aaaagcaaaa	22560
gtacaaaagta	gggtgtgatg	gcacatgcct	gcagtcccag	ctactcagga	ggctgaggca	22620
ggaagatcct	ttgagcccag	gagttggagg	ccaggagctg	ggcaatagtg	agaaaaaata	22680
aaattaaata	ataataataa	taaaataggc	tgggcacagc	ggctcatgcc	tgtaatccca	22740
acactttggg	aggctgaggt	gggaggatcg	cttgatccca	ggagttcaag	gccagcctgg	22800
gcagcaaagc	aagacaccca	tctcaacgac	aaattttaaa	aaatcagcca	ggcaggctgg	22860
gcatggtggc	tcacgcctgt	aatcccagca	ctttgggagg	ccgaggcagg	cagatcactt	22920
gaggtcagga	gttcgagacc	agcctggcca	acgtggcaaa	accctgtctc	tactaaaaat	22980

acaaaaatta	gctgggcatg	gtggcagatg	cctgtagtcc	cagctactga	ggcacaagaa	23040
tcgcttgaac	caggggtggca	gaagttacag	tgagccgaga	tcgtgccacc	gcactccatc	23100
ctgggctga	gtgagactcc	tgtctcaaaa	aaaaaaaaaa	aaaaaaaaaca	aggagccagg	23160
cacgggtggg	tgagggagg	cacagaagca	gcgcctcttc	tgggggcacc	cccaatctct	23220
agcgatccag	aggcctcagg	atcctgaagg	gagaaaaaac	gtgaagctcc	gtgctagaag	23280
agaccataga	gattggaatc	agctggttct	attttacaaa	aaaaggaaac	tgaggccctc	23340
agaaggtgag	tgccctctca	tgccccacag	ggaggcagg	agagggctct	gagccctgca	23400
gggcccctgga	ttcttgcaat	ggggtggagt	ggagcctgtg	ccgccccac	caggcacctt	23460
ctcaggagag	gagcctgtgt	catatccttg	aaggggtcct	tgagccctc	aaaaggctaa	23520
aaaccacttt	cctccttgag	tgaaccttca	cctcagttta	accacaagaa	aaactacatt	23580
aaggcccagc	gcagtggctc	atgtctgtaa	tcccagcact	ttgggagggt	gagggtgggtg	23640
gatcgcttga	gcccaggagt	tcaagaccag	cctgggcaac	atagtgaaac	cctgtctcta	23700
caaaaaacaa	caaaatcagc	tgggcgtgg	ggtgcacacc	tgaggctcca	actacttgcg	23760
ggctgaggtg	agaggattgc	ttcagcccag	gaggtagagg	ctgcagtaag	cgggtgactga	23820
atcactgcac	tccagcctca	gcaacagagc	agactcaaaa	aaaaaaaaaa	aaagcaggcc	23880
gggtgtggtg	gtcacgcct	gtaatcccag	caccttgga	ggccgagcgg	gaggatcagg	23940
agatggagac	catcctggct	aacacgggtg	aaccccgctc	ctactaaaaa	tgcaaaaaat	24000
tagccggg	tggtggcg	tgccctgtag	tccagctact	caggaggctg	aggcaggaga	24060
aaggcgtgag	cctgggagg	ggagcttgca	gtgagctgag	atcacaccgc	tgactccag	24120
cctgggcgac	agagcaagac	tccatctcaa	aaaaaaaaaa	attaaatctc	aaaaaaaatt	24180
acattaaggc	aaactaaaag	atgtttaaaa	tatatatatt	aaattaaata	cactccaata	24240
gagcaaatat	gaaaataccc	agaaaacaca	atccccgcac	ccccaggaca	acctcccagg	24300
gggtccacag	caagagaccc	caagcacgag	agacagagaa	cagtgtccct	gtggcggaac	24360
ctctggccca	tcaggctcta	ttagaaaata	aggctcttgc	cactgagaga	aagaggcaca	24420
gtcgcccagc	agccacgggc	tctggcacac	cacgagtcag	gccagcaaag	tgtcaactgc	24480
ccccatacag	gtgacaaact	aggacaaact	ggaaaccaga	ggctggacct	ggagcacagg	24540
gaccaccaca	tggggctggg	gaatgggcag	ggacctcaga	gcgccaccca	catgcctaag	24600
agcagcgcgt	atgcgcagtc	ctctgcagtg	cttagggaca	caggagctc	ccccacccc	24660
caaccagga	aggcagcccc	cactaccag	gtagggaaac	gataggacca	gcaccccggt	24720
ctgctcgtaa	ctcagggtc	caggccccc	cgggggcaac	cagcacagag	ctcagacccc	24780
aaatatcttc	acccacctcc	tggtccccc	ctggacaagg	gtgctgggga	ctggctctca	24840
gtcacaccct	cggggtactc	ttcaaaggac	agctggatgc	cccagggcag	gagcttttgg	24900
cccccagctc	cctcacccca	gacaccagct	cttgggaacc	caccagcatg	ggcaagggtg	24960
acaccatcgt	cccgattttg	cagatgagga	aactgagggt	gagggctggc	acacggctct	25020
ccagagctga	agagaatgca	gagagcagcc	ggagccagcc	ggtgggtccc	tgaggccggc	25080
tcgtagcaag	ccacagctgc	ctccgccc	cacacttgga	cctcactggc	cccaggacag	25140
ccctccagg	cggcctggca	cagagcccac	accttgcctg	ttcctgaaca	aataagtga	25200
caaggccacc	aagccgagga	cctggatgta	gccccggctc	ccgccagggc	ctccccaca	25260
gactccccat	ttggagagcg	cattaagtgt	ttccaaagcc	tcacaaacca	cagatgtccg	25320
gctgtctcac	ggcttctgta	acctgaactt	ggccctcact	ctgcccctcc	agcactcctc	25380
tcaggggcca	ggcccccctc	ctgagatgcc	agcactgact	ccccaaactt	tccccatcac	25440
ctggctcggt	cctgaacctc	ggcaggagag	tctcaggcca	gatcctccca	ccagccacct	25500
ccaccaggat	gcaggaggca	tgagacctgc	tcgtgccggc	tgggagatgc	aaccaacca	25560
gatcaatcca	atcagcggat	gaactgacaa	atataatgtg	gtccctccac	acaatggaat	25620
attattcagc	cacaaaaagg	gctgaaatag	gcccggcggtg	atggctcaca	cctgtaatcc	25680
cagcactttg	ggaggccgag	gcccggcagct	cacttgagg	caggagtcca	agaccagcct	25740
ggccaacatg	gtgaaatccc	gtctctacta	aaaatacaaa	aattagctgg	gcgtgggtggc	25800
gggcacctgt	aatgcaagct	acttgggagc	ctgaggcagg	agaatcactt	aaaccagga	25860
ggcagaagtt	gcagtgcagc	aagatgcac	caccgcactc	caacctgggc	aacagagcaa	25920
gactccattt	caaaaaaaaa	ataaaaggct	gaaacaccca	tacgtggtac	tacttggatg	25980
actcctgaaa	acgttacagt	aaccaaggaa	gtcagccacg	aagacgcatt	gtaagattcc	26040
cttcacgcaa	aatgcccaga	acaggcagaa	ccacagaggc	agaaagtcca	ctgggtgttca	26100
ccaggggatc	cggggagagg	gaacgggaag	tcaccgtgta	atgggtatgg	gttttatatt	26160
gggggtgatg	aaatctctta	taacttgata	gaagagagg	ttgtaaacac	tgtgaatgta	26220

ccaaatgcct	gccttctata	ctttaatatt	ttatattata	taagtttcac	ctcaatttaa	26280
aaaaaaaaa	actcgacacc	tttcacctag	gaaagatctg	gcttttagctt	gcatttccctg	26340
taactcctgc	ctaaagcctt	ccagaagctt	ccgctgcctt	gtggatcaca	accagactcc	26400
acaccatgat	ctggcctcta	agggcctctc	gcaggacacc	ccgaggggtga	aggagcacc	26460
gtggggccac	ctctgcatag	ctgcaaagct	tctttccctg	tcctcccctc	tacatgggaa	26520
gctctgccc	caggggcggg	gccttatctg	ccattctatc	gcactcaacc	ctagcacttc	26580
actcggtagc	agacaccaa	gcaaacagc	aacagcatta	taccgggcca	ggtgcacggt	26640
aactcactga	attcatggta	ggaaggattc	tattcccatt	ttacaggtga	gaaaactgag	26700
gcacacaaag	gtagcatcag	cttcctaagc	ctcccagcac	aggaagcggc	caggctggaa	26760
tcagaccctg	ggcgcagggg	ctctgtccac	agtgtact	aactactcct	gcccccgagg	26820
gctgcagcgg	tgagttagtg	agtttgtcag	tggactggat	gtccaagggtc	atacaggaaa	26880
aatccagact	attgtaataa	cagcctctag	accggctggg	gccagaaaga	tcgaggacgc	26940
tgacacacaa	ctgcgctcac	tgcagctctg	ccagggatgg	ggctaaaggt	ctcacacagg	27000
gcagttagg	ctccccatag	cctgggagag	gaacgggggtg	agataacaga	aactaggtat	27060
ggtgcccgaa	gtcaaacagc	cactgagcat	gtaaaccacag	gtgggtctga	cccaaaccc	27120
ctccaccccc	atcagccctg	caaccctctg	ctgcaaggga	gaaagcaact	cagaggcctc	27180
acctgcctac	atccccacc	cgtgtgtgtg	agttctacta	aatgcctgag	cagtgcacaca	27240
gcacggctga	aattaaacgg	gttccaaaaa	cgacaggaag	cacgaagtga	atctcccag	27300
gaaagtgtcg	aacaaatgct	ggatcgggtt	caccggcgaa	tttcttgga	ctgaagagg	27360
gagctaaaca	cacggggccc	tgctttggag	gggactctct	cagggtgtc	cacacagcac	27420
ttgggttaacc	ccactgagcc	cttctgggct	ctcccagagg	gcccggcctt	ggccttgggc	27480
atctacagga	ggaacctcca	gggggagagg	gggtgcctgg	acaggccggc	cctggaacaa	27540
gcacttgggc	cccagggaga	gaggactagg	gcttgggagc	tggggaagtt	ctcagcactg	27600
ggaccactag	aacaaagcca	tttcctgtcg	ttcacagctt	ccaattgcaa	caggaagcaa	27660
tcaggaaaaa	taatttagcg	cccacttact	ggcttcgctg	aggtccgagg	catgtatttc	27720
acacagtaaa	accaggggata	taacatcaaa	accgttctgc	agaaagattc	ctccctttcc	27780
ttccattttta	ggcctggatc	accacattca	ctggggctcc	caggccttgc	tgcctaattgt	27840
taaaataatc	aactctattt	ttgcctcaca	cacaactgaa	ctctacagct	ataattcttt	27900
ctcctcaggg	gctcgaacca	catggacgac	aggcatttga	ctccagcaac	atcaccccaa	27960
aacgtgcaca	aaacccaaaa	ctgcaatgag	gtgaaaggca	acgcggtcgg	cctagaaacc	28020
ccccctttta	aacaaacagt	ttccccaaaa	ccccctttgc	ctccttgacc	caggcatttc	28080
cggaaaaagg	agcggcgctg	gcctgtactc	cccagatact	gtcgctgttt	tgtcttcacc	28140
ttgttttgc	agctccagac	aaggccccac	aatgtaaaca	cgctcctgaa	agaggcagat	28200
ttgggggtgaa	actgtccata	gaatctctag	gcttgggtca	gaggcaggag	gacgtgaaac	28260
aaactccaag	ctcctcctgt	tccccgctgt	ccccacacc	tccaagcaga	ggctgcagcc	28320
tgggggatct	gactacagg	ccaccccgct	gcaccattca	cactggaaat	attcaggggag	28380
acagctgttt	gccttaagga	ggcccagaca	aaggggccc	aggtcctccc	cgctaaactg	28440
ccacaaacag	aacaggagcc	gcggcgtgca	caggcacttg	cggccgtgcc	acttggccag	28500
ccatactcca	gaaaaacaaa	acacgcacat	ccgaagagaa	tgatttaggt	agcaagaggc	28560
ttgcttgaaa	aaccacatgg	caatctccaa	attaaagaaa	catgtgtagc	gtttcacgac	28620
tgcttaagtt	tcctgagtcc	tcctgacctc	aactccacc	cctgggaaac	accaaagtt	28680
ggatgagaaa	gttccccgc	cctacctctc	cccacgggag	tgtacaactg	aggcacaagc	28740
ctgcctcccc	cactgccccg	cgatctggga	ccacgtctcc	tccgcgtagc	cgacccgggg	28800
atggacacta	tctggggacc	cggcggccac	acggggcatt	cgggtcgccc	gggcacctgg	28860
caggtgtcag	tccgcttgga	aaccacagc	cacgcggctc	acaggagcag	cgccaccggc	28920
taggccgccc	cgcgcccggg	ctcagaactt	tctcgctgcc	acttcagccc	gtcctcgag	28980
cacgcggggc	ggccgcgcgg	ccgtgggaaa	caggcttgcg	aaccggctcc	ccgggcccag	29040
ccgcctccg	cgcctcaagt	cccgctcgg	tggccggccc	gggccacacg	ggcccagcgc	29100
gggctcggct	cggctcccgg	cttccgcgg	gctcgggcag	gtgaggaccc	gcccgcgcgg	29160
cacctggcgg	agcgggcggc	ctcctcgcca	gcccgggacg	cagcgtcccc	ggggaggggc	29220
cgggtgggga	gacaaagggc	ccgcgcgtgg	cggggacgcc	ggggacggca	gggggatccc	29280
gggcgcgcgc	cccaactcgc	tcccaactcg	ccaagtcgct	tccgagacgg	cggcggcgcc	29340
cgcgcacttg	gcccgggggc	cgcgggggc	attgtccgag	caaccgcggg	cccgtcttac	29400
acgccggggc	cgggaaggta	tcgaatcagg				29430

<210> 8

<211> 33769

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33739), (33749), (33758)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 8

```

cttccccctta cactggctcct tcgacccgcc tcggatgaaa actgaatggg tttagcctta 60
gaggtctctcg gtctctaagg gaggtgggtc aggatgccgg ggacagggtc ctcttcctcg 120
ggcaacgtgg gggaacgagc cacctacccc tccactgaat tgccctgggg tgtgggtacc 180
gacggctcat tcggtgtcca ggggtctgaga tgtgttgaca ggaagaatga aaggggatgg 240
gagggatggg gcgaaagaag ccacctgcag cccaggaac tatctggcca gcacaccgtc 300
acccagcggc ctgagccacc cctgccagag ccaggaggag accctgcca tgggtcacca 360
gtgtgcagga actcagaagg tcatcacagt taataccctc catgccccaa tgtgggaaaa 420
caggtttttt cacaacaaac aagataattt ttgttatttt ggcaaaagga ggcaggggcag 480
ccccggacac ctccatccca cctcatcacc cagccgcagg gccccggcca tccctgcaga 540
cagagtggat gtcacaacct ccctgcaccg aaccaagtgc agctcccagg ccacaggcca 600
cccaggaaag gtccagtggc ccccgagggc tcccaccgca ggccctccac cacagcgggc 660
accaaccag gatagctgtg ttctcctggc ttcttttcac acgggtagca gaaagctgag 720
atccggggaa agctgagatc cagggaagc tgagaatcgg cctctgctgc ccggacgccc 780
acccccagct ctgctcccag ctccaggggc tctttctcag gtgcccttac aggaggcaga 840
gggcttgagc cacctcctgg gcctggggca cgcaggatga acggggtcac ggtgcaggcc 900
actgtccact gcgcagatcc caaggccata aacagcctgg ccacagtggc tcccagctg 960
gcaggcggcc agattatttt tgttgttttag caattgatta agtttctcag ctgccccag 1020
gggtaagtgg tggggcaaat gccgcaaccg cagcatttga ccggggatcc tgtgccaagt 1080
gaccataggg tcacaaagca caagggaagt ggctgggccc gatgctggct ctgctggaac 1140
ctgaggccgg ccactgtcac ctgcacggtg cctgggaact tccagcaagc acagagaagc 1200
tatggccctc caggagcagc tggcaggcac cttggcctgc agtcaggggc tctgtctgct 1260
cagctctaaa acaggaaagt cgctgctctg cctgggggtca gggcagccag agagtgaaca 1320
agtcagtgcc ggccctcagga agggacctgc aggcgggtcc ctctctctcc catccctcgg 1380
tgccagccag cccctcctgt ggccccccac tgccctgctc tgccccatg cccaccaca 1440
acctcaggcc catggctgca tggccactcc ccaggcaggc agtggggatg ggatttcacc 1500
atgttgacca ggctgggtct gaactcctga cctcagggtg ggagttccta aagtgtctgg 1560
attacaggcg tgagccaccg cgcagccct cctgttggtg ctaaacactc acacccctt 1620
gctggggacc ctgggtgagg aacacagcct cacaagtga gtgtggtttt gttgagcaaa 1680
tgacgcctgg gcagccctct catctttgcc taaaactgaa gaatttaggg gcgtggatgt 1740
ataaaacagt tggtgactta aatgaaaaag aaggccacac tcccccttt aggcaggcgg 1800
cctaattctt taaaagccag cacagggtgc ctttctgaac ccaggcacac agtaggtgtt 1860
caatggacag cagcggttac ttgtactgct catgacacc tgtctgtggc ctctgcagct 1920
ggctccagcc tgacgcatgg ctgcgcccct ccgcaaggcc accccggtat acatggaaac 1980
tctgtggaga aggccttggg ggccggccag gacgccaggc ccagatccca tctgcgcct 2040
tctccatag acctcagcga gctctcggca ccattgtgct caggccatt taagaagtag 2100
ggccggccag gcatggtggc tcatgcctgt aatcccagca ctttgggagg cccaaggtgg 2160
gtggatcacg agatggtcag gagatcgaga ccactcctgg taacacggtg aaaccccatc 2220
tctactaaaa atacaaaaaa taagccaggt gtgggtggcg gtgcctatag tccaagctac 2280
tcgggagggt gaggcaggat aatcgcttga gctcagcagg cagaggttgc agtttagcgg 2340
gatcgcgcca ttgcactcca gcctaggtga cagagagaga ctctgtctca attaaaaaaa 2400
aaaaaaataa aaaaaagaag cagggccagc cacggacgac ccctcacaca gctcccagga 2460
cgcgctgcct ggtatagggc tcaggaccat gaccgctgca gtggcccca agaaacgtta 2520
cttttgtcac ccaccccgcc tcagtggcag tagccaaaat aacggattag aatggaacca 2580

```

tgtgacaatg	ccactgcccc	aactgacaga	agatggctat	cagcagttca	cgcgccccca	2640
cctatcacaa	gtgcagggca	ctctacaact	tatgcaccc	tccccagaca	ccgtcccttc	2700
gacctccca	ggtcagcaag	gcacacaggg	cctacatttc	acagccacac	agcagagggc	2760
tgaggctgga	actcggatgc	tctgatttcc	gttcaatcac	atccccagag	gtggccacaga	2820
gacggggggc	ttctcttgac	aaagtcaaga	aagtcactgc	cagctccact	gaagaccaa	2880
gaacctcagc	tctcaaaacc	tcttgaagg	gttaccgaac	tctcccagcc	tgtttcctgg	2940
gtcccgatgt	tggtcccg	ggacacagga	agaggaagaa	gctccctaga	gcagagcctg	3000
gtgcacctgc	cacactctca	gagggctgcg	cacgggcgga	ggagccgtgt	gcaggagtgg	3060
ggctctggatg	gaggggcgct	gtggccgggg	gcagggggca	ggggaagggt	gctccaggtg	3120
gtgggcacag	cacgagcagg	ggcagggagg	tccacactca	gatgtgcaca	gggagaaaca	3180
aatcgtgcat	ttccattgga	ataggcggta	aaaggtagaa	aaacagagt	ggggccagga	3240
agggagtcgg	agccttctag	tgtctctctg	caggtgagcg	gcagcccag	gtgtcagctc	3300
agcagacttg	gggtccaggg	gccgtgtctt	ctatcactga	ccccaggga	cacggaactg	3360
gggagggaga	gcagaggcac	agggcacggt	cagtgaacg	aaacaaggag	tcatcaccaa	3420
atgcggaaag	ggcaaggagt	gcccgcagcc	gcacaagggt	tctgtctggg	caacgtgggc	3480
gtcccaccag	gccccgcacc	ctgcaagcgc	aaagctcgcc	actgaagata	aagggaaagt	3540
gttgagagctg	cggagctgg	ctggggctcg	catggagctg	ggcttatgct	gcagtcacaa	3600
gggggacatg	gaagaggctg	caggggacaa	aaccagtgc	cacagtctaa	ctctgagcct	3660
gtggaaaggc	gcccacagca	ttcacccatc	ccagagatgc	cattccccct	gtgccccgc	3720
tccacggtga	cagcgttctc	caggaatatg	atgcgcccct	ctcctcttgc	atcagccctg	3780
acagtga	ttcaggccaa	aaagcagaag	agcacagctg	cgtggttcca	tttccatgta	3840
gttctggaac	aggcaacgct	aatccaaggt	gatagaagtc	aggagagtgg	tggagggggc	3900
gggggttgag	gatggcaag	gggcaccggg	aactttccca	gtggtagaaa	tgttctctgt	3960
ctggaccgtg	tggtagtatt	gcagacatat	gcagctgtca	aagttaatcc	aatgtacac	4020
gttaaaatgt	gtgcgtttta	ttgcctgcaa	gttatacctc	aattaaaaaa	ataaagttag	4080
cactcaggct	tcttccacaa	cttctgaac	cgtgtgagct	gattttcttg	ctattaaaaa	4140
ttcacggtcc	atggctgaga	acagcagctg	ccttctgttt	gcaaagtcaa	cgccaatcac	4200
tgcccggccg	cggcagactc	ggccccacag	gacctccttt	cttttttccc	tttgacctac	4260
ttccctgata	agtgacaaga	cagccagact	ctgggaacaa	acgcccgtta	ttcgcccccg	4320
agctgagcgg	gccctgcttc	ctgagcta	ccgcccggac	agacggaggg	acgtgagggg	4380
ctttgccgtc	ggctccagct	gtcagctctg	ccgtcagact	cgacagtggc	cccctctgtt	4440
cctcccgtg	ccccactcc	atccccgact	tctttttgtt	tctgtccct	gacagacgaa	4500
catctgttaa	aactctgtct	gggtgagctg	tggccagcgg	cccacaaatc	cccaagccgc	4560
accccagcct	catctggg	ctgccgggag	cactgcctgg	ccacctctg	gacatagctc	4620
tgagagccac	cggccagggc	acgtgtggcc	cgagtggcat	ggtgcacgcc	gctaagccca	4680
ctgccc aaag	gccccaaagc	aggagggatg	tgcaggagac	aaaagtcaaa	agaacagggg	4740
caggttccac	agaggatggg	gctggagggg	tggcagtgcg	gaacagcagc	ttccgaggat	4800
ggcgggtggca	actcccaaat	aaggcctcac	tctgtctgtt	tttagctcat	tccacataat	4860
tggaaaaaca	tggcagaaac	cgaagccagc	tgctgccttg	gtcctggggc	tgtgtggagg	4920
gggtggggag	gccggaggcc	caggctctgc	actgcactgc	tggggatgag	agtgcactctg	4980
agctgcagag	agcagcatcg	cagccgccat	ggtccattg	agccccggcc	acgtggggcg	5040
gcagaggctc	gtgggatata	cctgccctgt	ctcatggggg	tacttccagg	aggggcgggg	5100
gagccaggac	acagcccagg	gctagcggtc	accctgcagc	tcagggggcca	cgtaaatagt	5160
gccaccttga	aggcacacag	cagtgcgggg	cccccccgcc	caccaacgca	tccctacctc	5220
taggaggccg	cctgtgtgcc	cctgggaacg	ctgctccctg	tcccttgggg	tccctgggtg	5280
accaccctct	cagccccttc	cttggggaag	gcacctgact	ccctacaccc	agctggcttt	5340
catttgctca	aaatcaggaa	aaagcagaat	tcaagacatc	acagaaatgt	cttcgcctgt	5400
aactccatga	aagataaacg	gtcagacacc	caggagggag	tcccagggac	ccttgagtct	5460
cacctgaggc	tctggcttca	aacctcgaga	tgtttccagc	catgctagcg	ccgcccccca	5520
caacctgccc	cacacagtcc	tcccttggga	actcacagat	ttggccccca	cctgccccgt	5580
ttcttctggt	ggagtgggtg	cgttgggttg	gggtggggct	ggggactctg	gatgtgtctt	5640
aagagtctga	gtgattctga	cacagccagg	ccctgcccc	ctcctgacct	tgcccccaca	5700
ggaaagggag	ccacacgcct	gaagcgccca	gcacaccccc	ctccgtcctc	cccaggctac	5760
ccgtggcg	tgtgagccgt	gctccccact	gccccctcac	ccaccccgag	tccctcctggc	5820

agcaccagc	cttggagct	acttctgatt	acaaccgccg	aaggaagact	cgctccctcg	5880
gcactgaccc	agacagcctg	caccatcacg	ctgctcagca	caaccacac	agccttcctc	5940
caaaccocat	ggagcgggga	gtataatcac	cccctttcta	ccaacggaca	aactgaagca	6000
cagagagggt	aagtcacttt	cctaagctcc	caacacgatg	acaaaaata	gaaggtcagc	6060
ccgcaagtgg	aactaggtgc	tccaagtccc	cggtctgcct	gacactgcac	ctcctcgccg	6120
ccacggtccc	gggtccgcct	gacactgcac	ctcctcgccg	ccacggtccc	gggtccgcct	6180
gacactgcac	ctcctcgccg	ccacggtccc	gggtccgcct	gacactgcac	ctcctcgccg	6240
ccacggtccc	gggtccgcct	gacactgcac	ctcctcgccg	ccacggtccc	gggtccgcct	6300
gacactgcac	ctcctcgccg	ccacggtccc	gggtccgcct	gacactgcac	ctcctcgccg	6360
ccacggtccc	gggtctgcct	gacactgcac	ctcctcaaca	ccaccacggt	ccggggtctg	6420
cctgacactg	cacctcctca	ccaccaccac	agtcccggtt	ctgcctgaca	ctgcatttcc	6480
tcatcaccac	agtcccggtt	ctgcctgaca	ctgcatttcc	tcatcaccac	gggtcccggtt	6540
ctgcctgaca	ctgcacctcc	tcaccgccac	gggtcccggtt	ctgcctgaca	ctgcactttc	6600
tcaacaccac	tcttggcccg	gctcccaact	acaaaccaag	ccatgtcttc	catcctgaat	6660
cctcttggcc	taaacatcac	tcacaatgcc	tccctcgggg	acaggcaca	gtcccaccag	6720
cacagcctcc	ttcggtacct	gcggttccgc	tagcccgagg	ccagctccag	agccctcacc	6780
acagagcctc	tatccttcac	ccccggacac	tggacctcac	caaccatag	cctggaggag	6840
atccctgtgt	gaccccgagg	cctcctctgc	ccgactctga	atttcactgc	ccaacgtgac	6900
acctcggaag	gctctctggg	cactggcagc	cctccatggg	caccgctcct	tctggccagc	6960
tctgacatcc	cggtctgtga	ggtgccctgc	acgaggcctc	tgcccactgg	gacctcacag	7020
ccgtgctgtc	agctgcaaca	agcgacagaa	tttcacggtt	tcttcacggt	gcccctgggt	7080
gagcagctcc	aggtagtttt	cagtcgaggg	gaggcgctcc	gtcagcagcc	aggcggcaca	7140
gctaattcat	gcccgcgggg	cgcacggccg	caataccaat	gggcacctgc	agcctggaaa	7200
gccacagagg	aaccgagaac	agcgactgtg	ctcaggtgac	aggactgtgg	tcttttaaca	7260
aaacattttc	ctttaacgtg	atattttacg	gcaagggaatg	aaacctggag	ggcaggacat	7320
ttggatacta	aagccccagg	ctgccgcgtg	gtctgctttg	tgaagtctga	agcccgccgc	7380
ccattctggc	cccgtcacca	ggtccggctc	tgactcacca	gcttcaatgc	taggcccgtg	7440
ctgtcctcca	accagaacat	gacttcctta	aggacaaagc	cgtttctcgc	ccatccccat	7500
ctccctctgg	attaagaaat	atgggaagat	cttctagaac	cacctcaaat	ttgcagagag	7560
ccatcctggg	gacaaaccct	tgaaatgctt	ctaagaagag	tttaggtttc	ttctcaactc	7620
taaaacctct	agaaaactct	atttcacac	cagctgcccc	tggaaacttt	cagcttcaaa	7680
agggcccgag	gcaggagagc	ggaggagcca	gcatccacac	cgagcaccag	cctgttaatt	7740
aacgggaagc	gggtggggcc	catctccagg	cagctctgag	gtcagactgg	ggaacctatg	7800
ttacaaaaaa	aagtgaactg	aaacgctcac	gtcctcatgc	aaaaccagac	tcccagttgc	7860
atctttctgt	ctcattgagg	agctttttcc	tccctttgac	agaacaccct	acacacggca	7920
tctggaacca	aagcagaaag	attcaggctc	agagtaaaac	agtccccaca	ctggctgcat	7980
gtggacgttc	ccggcccaga	gtctcgccca	agcagggcct	ataaatgaca	caaaatgttt	8040
ttctcctgcg	tgccagtcac	gctccaactg	agttatgtgt	aaaagtgcct	ctcacggctg	8100
agggcaaaaa	cagttccac	aagactagag	aaagggtgacc	cctgacggct	gagtctctag	8160
ggagcgtgga	gctgcgtgct	cagccctgcg	gcctgacggg	ctctggaatg	gaaaagctat	8220
ccaactggaa	gggcagggct	cgctgctagt	ccagcgggtc	aacccacag	gtgtctgtgg	8280
tgtcagctcc	atgccacaga	gcccagggct	ggggccagag	ccaccaggcc	ccctgccagc	8340
ctgcaggggc	ctcctcctct	gggtagccta	accacccct	gtgagcgag	gcagcctcct	8400
ctaataacca	cagggcctgt	ccccccctct	cccccgcttg	caggaaaatg	agccctgagg	8460
actccccagg	gctgctctgg	gcctggacat	ggagactggg	aattacattt	gcagaaggag	8520
cgcaatgccc	ttgaagggct	cagccacgag	cagccagtc	ccagggtcca	gaaggcccag	8580
ctgttagaac	cctgggagcc	agcaaagagc	caggggctcc	acctaagtct	atagcccttg	8640
cctcttctgg	ttgggaaaga	aatcaacgcc	cctttactgg	ctccactga	cagccactc	8700
ccccaggtat	gggaggattc	tgggacgatg	caggcaaacc	tggaccctga	gtgaacctgc	8760
cccagctctc	acgggcctgg	caccagccac	agcacctaag	gcgcgggtca	tggtgacaac	8820
atgaagggtg	taagggcctg	gacagtggac	atggcagctg	gacactgggc	accactgga	8880
tgccaggcac	ccagcacggc	tccgtcacc	ctggatgagc	agtggccctt	tgcaagccag	8940
ggtagcctgg	gcaagttatt	tgggggtctc	caagcttgct	cagctgtgcg	acttcactga	9000
gccatgagtc	tgggatttta	tcaggggcca	caccgcgttc	tggaaactctg	atacgtgagg	9060

gagccacaca	gggaccctta	acaaaagctc	ccagggcaac	atgttctctt	gcctcagtct	9120
cccaaatagc	tgggattaca	ggcgacacac	taccgcccgg	ctaatttttg	tatttttagt	9180
agagacaggg	tttcaccatg	ttggccaggc	tggctctgaa	cccctgacct	caaagtatcc	9240
ttccactgtt	agggcaaggc	acctgacagg	cacgactgca	cgatctgctt	gttgggggct	9300
gtgtccattc	cccactcctt	cgacaaatgt	ccacaccag	ccttgctttg	acaccccaag	9360
aacagagatg	gtgacacctg	cttcctacat	gcccattgct	ctcccaaggc	agacatcccc	9420
agcagatgca	acacagtgtt	taggcagaca	tcaccaatcg	atgggtggca	cagacaccag	9480
gccctgctcc	ctctaactcc	agtggccagg	ccccaagcca	gctctcacct	gcccactccc	9540
aaccacacagc	agcaagactc	agaaatggca	aaaacacaaa	gagaacagaa	acgccccata	9600
gcgggaggat	gactaaaaga	catgtcttga	taagatattg	ttcaggcata	ggccaggcac	9660
agtggctcat	gcctgtgatc	ctagaacttt	aggaggctga	ggtaggtgga	tcacctgagg	9720
ttaggagttc	aagaccagcc	tagccaacat	ggtgaaaccc	catctctact	aaacatacaa	9780
aaattagcca	gacatagtag	cgggcgccctg	taatcccagc	tgcttgggag	gctgaggcag	9840
gagaattgct	tgaacctggg	aggtggaagc	tgctgtgagc	cactgtactc	caacctggac	9900
aacagagcaa	gactctgtct	caaaaaaaaa	aaaaaaaaaa	gatatccttc	actaaaactc	9960
atgtctttga	tacatattta	cctcctgcaa	tcgcaaatgc	ttctgcagtg	cataaagtga	10020
aataaatagc	aggaagcctt	acggttcgat	caccacacac	gacacacagt	cacatacagg	10080
aaaaacgcag	ggagggtctg	ggaacaaaaa	aacagaagat	aaaatgtgga	gacagacaca	10140
ccaagagagt	aagagaccac	ctccagacct	cccttcagct	tctcaaacac	acgagccggg	10200
cccgttacag	aatttgcg	gaccgctgca	aatgggaagt	gcagacagcc	ccttactcaa	10260
aaggtaggaa	tttcaggtca	acaacagagc	tcacctcata	tgactacaca	ggtcacacag	10320
cccgtgaagt	cgggtcccaac	accagcatgc	tcctgcctca	aagccgctgc	acgtgctgtt	10380
ccttctcgcc	tttccctctt	ttagtccctc	agatctcagg	cctcctgaga	gagacctctg	10440
acctgccggc	tcaggcgcc	acacccccag	tacaggagtc	tccggctcag	cccctgctgt	10500
gttccgtacc	cgatccaggt	ctgtcctatg	tccatctgtg	tgccggcttg	cttccctgaca	10560
tggcccccac	cacacgtgtg	cctcggggca	ggggaacagg	cccgtctcat	taactgcttt	10620
cttctcagat	attttctgga	atatttgtgg	atattgggca	acatatatgc	tccacctttt	10680
tcagactagc	caggacgagc	tgcatTTTTT	TTTTTTTTTT	tttgagacag	ggtctcactc	10740
tggtgcccg	gctggagtat	agcggcatga	tcttggctca	gtgcaacctc	cgctccttag	10800
gctcaagcaa	ttctcctgcc	tcagtctccc	aagtagctgg	gattacaggc	ccgtgccact	10860
actgccagc	taatttttat	atttttagta	gagatggagt	ttcaccatgt	tggccaggct	10920
ggtcttgaac	tctgacctc	aatgatcca	cctgccttgg	actcccaa	tggtgggatt	10980
acaggcgtga	gccactgcgc	ccggcccag	ctgcctgttt	tacacctttg	ccatattccg	11040
gtgattctct	ctcccctccg	tccccggcc	ctgactgtgg	tggccactcc	ctgcgcgtcat	11100
gagcccgat	gtcctcactc	tttcccttcc	cgccaggact	tcaaccaaca	ctgcagagcg	11160
cagggtccag	ctccagcact	gagttcagcc	tcttctcacc	aacagacagg	caggaaagaa	11220
aacaaactct	gagaaggcca	aggttcccgg	gcagccagca	agccaagcat	ccttctccgc	11280
tgaggcttgt	gcaggcgagg	cacccctccc	tccagggagc	aggcagcgtc	ctggggcagt	11340
ctgcgaggga	gaccagggcc	cctgctccac	cagggcccca	ggtagggggg	cagcagcaaa	11400
ctcatggctc	tgggagccag	accccacctg	ctagaacctta	ctatgccacc	tgctgtgggc	11460
aacccagggc	tggtgacttg	ccctggcctc	ctctgtaaac	aaagggtcca	tccaacctgg	11520
tcaaaccact	cctccccttc	aagggtctat	aatcctccct	taacctgctt	gggtccaaacc	11580
cctgggtgtcg	ccaggtcact	caggaggcag	ctcatctgga	ctccttccct	gggtccagtt	11640
tctctctcaa	cattgccttt	gaggccgagg	tgaacggtca	acagcgaagg	gccccagagg	11700
tgatggagga	gcgggtgtcc	aagacactca	ccctttctaa	tgactgact	ccctcgtgga	11760
ctcacttggt	ccgtctcccc	caccaccca	gccccagagc	ccagagtgcg	agcgccagag	11820
gccccgggatt	ctgtctgcac	cgcggggtcc	ccagtgcctc	ggagcaatgc	cagcaccggg	11880
caagtgttcg	acaaatgcct	gctgaatgag	caaattgatg	gatgaacgaa	tgaatgagca	11940
agcagatgaa	tgaatggggt	gctgtccaga	gccgtgagga	ctaggccggc	caagtcccca	12000
tttctcaa	tctccttctc	ccgacttggg	aaacaagatg	ccttggtcggg	gaggctctcc	12060
aaccatcccc	tgcagcagcc	ggcacagcgg	acagaccctt	tgatgtaaca	gccatgtctt	12120
cattaaagat	gccctgctct	cagaaaagaga	aagacaaata	caaacctgga	aaatcctcac	12180
caaacgcagg	acccctgcca	gggagcagag	aaaagaccca	cacgccacgg	gcgccacgac	12240
cacacacaca	ccccagccgc	tgacacacaa	cacagaccct	agccagcaag	aacaggggga	12300

ccaggaaact	gttcctaaag	tcaggacccc	catgtgctca	gacagcagtg	agagcaagga	12360
cacttctcca	tccaccggat	gccaggagag	tccttttagg	gggccccaca	ccgagactct	12420
gcccttagga	ctgttctga	gtgtggaagc	cagcccactt	ggaagccccc	tgccctcccg	12480
agtgggacac	cggcacagga	agcaggccct	gtccccacc	actttctgca	agctgggccc	12540
catcacgcta	cagaaacggg	gaggactggt	cccagggatg	gcgctttcct	gacacctctc	12600
gttaccacct	cgcttgccag	gccccagggt	cagccccaga	ggccagactg	gctatcccag	12660
gccccgggagc	atccccgaag	gcgagctgca	tcctgaacgt	gtgtgatttc	ccgaagggcc	12720
cgccccgaac	cgacacctgg	aaagaaagat	cctcagccgg	tgccccagag	gagaagagcc	12780
atgcctcact	gcaacacagt	cccaggaagc	accaagtgcc	tgaggaccaa	ggcggagagt	12840
aaaaaagtgg	aaaatatctg	gggcaaaaat	aaaacaaaac	aaaacaggat	tgacctcctg	12900
ggctcaagca	atcctcccaa	ctcagcttcc	cgagtagctg	ggaccacaga	cttgaatcac	12960
cacaccgcgc	aagtggatca	tttcgaacgg	gtttgcccag	gttccttctg	gggcaccccc	13020
ggcggcgcga	acccattccc	gccaggcccc	gccccgcccc	cccgccccgt	cccgctccac	13080
cgctcacct	gccttacacg	tcctgccgtt	gtcctgcagc	tgacaccccg	tggggcaggc	13140
gcatgtgtag	aaaggctcgc	ttggggacag	caggcacagg	tgaggagcagc	cgccattgtc	13200
ctcctcacag	cgagtgtgga	ctgagaaaac	caggacagac	tgagagaagg	ttccagaaga	13260
ggaccgtcac	ttgtttctga	atgagtcaca	tcctgcctcg	tcccccgtag	cagcctccag	13320
tgtgtccctc	tgcccaaaca	tcggcctcaa	gtggcatcag	ggacctcccc	gcgggcacca	13380
ttccacctgc	ctcatcgctg	gccccgtcca	catggggccc	tcagcctggc	cagacggcct	13440
gcaatttccc	caaaaccagc	cgtagacctc	ctggccaccc	tcacacccag	atgtgacctg	13500
cccatggagt	gacatcctcc	ccatctgctt	cctcccacca	agctcctatg	actagaacac	13560
cctccccagc	tcctcggagc	cccaaaggga	caccctctg	caaaggctgc	ccccacgct	13620
ccaatggcgc	gggtcaggac	ctgcctgtgt	ggtagtgacg	ggaacccccag	agacaatggg	13680
ctcctgggca	aaaggcttgt	cttgtctttg	tgctatgtgt	ggaccacagca	gcttccatag	13740
gaacactgtc	cttcttgctg	ggatggccaa	gcttgtcact	ctcccaagcc	ctcctatgac	13800
caacagcaat	tgaacggaac	tcgataaatg	cttcacgac	ctcattcaaa	ccaggggaaa	13860
gctgggtgta	gcagccccaa	aatacggata	taactggaac	aacaaaactca	tcaaaatgaa	13920
cctctccctc	cctcatgctg	ccccaaagtgt	agatgggttt	tgtgaccacg	actttctcac	13980
caggaaacag	ctccagagag	ccccaccctc	ctgtgtcctg	ctctgggaac	agctggcacc	14040
cctaggcccc	acatttcaat	tcaaagtcca	aaccttccat	aatggcctgg	ccagaaatct	14100
ccatccctgg	tcctgtggg	agtgggccac	tgtccccaga	gccgcagccc	cactgtcaca	14160
gaagctggtg	catttcccca	tcagggacct	ctgtcacaa	ccagcgtggc	ccccaggctg	14220
agaactgctg	attctgggca	gattattcat	tgataaaatc	gcgacttgca	gggccaagca	14280
tggtggctca	tacctgtgac	cccagcactt	tgggaagtca	agggtgtgag	atcactggag	14340
cccacgagtt	tgagacaagc	ctgggcaacg	tggcaaaatc	tctcatctct	attaaaaata	14400
catacacaca	cacacacaca	cacacacaca	cacatatata	tgtatatata	aataaccata	14460
tatatatata	cacacatacg	tgtatgtgta	tataaatata	tatacacaca	cacacagaca	14520
acttcttctg	ggccttgaaa	acgaggcaac	cttccttgga	aatccccttg	ccactgctga	14580
gcctgaaata	gcccccatga	gctctgcaga	gggtcctct	gcaggccccgt	gtcccccagc	14640
cagccacaca	cctccctcca	ttgcagcagg	tacccttta	gagagggggc	ccccagagc	14700
atgggtctct	gcagggagg	gtcacctgcc	ccccacccc	accacgccc	gcgcacccc	14760
acgccccgc	atcctcccac	tcccctgccc	cgcgcccccg	ctccccccag	ccccctcacc	14820
ctctccccgc	tgcccaacc	ggcactcaca	aaaaggctgc	cgctcctggc	tcagcacctg	14880
gatgtccatg	ggtagtata	gggcactcag	gatctccttc	ctcttcccc	cagtgcgctt	14940
gttgaggca	tggatggagc	gggtctgcc	gtctgtccag	tacagagtgt	ccccggagag	15000
cgtcaggcg	aaggggtgcg	tcaggctgcc	ctccaccacc	ttctgcctgc	agtcagggaa	15060
gcgggggtgga	ggagccatca	ggagggtccc	ccgacagtca	ttgtgtctga	cccaattaat	15120
ttcttttttt	ttttttgaga	tggagtctcg	gtctgtcgcc	caggctggag	tgcaagtgat	15180
taatctcagc	tcactgcaac	ctccgcctcc	cgggttcaag	caattatcct	gcctcagcct	15240
cccgagtagc	tgggatcact	gatgccacc	actacgcca	gatgatTTTT	gtatttttag	15300
tagagacagg	gtttcatcat	gttggcaagg	ctggctcga	actcctgacc	tcagggtgat	15360
caccacctc	agcctctcaa	agcgtggga	ttacaggcgt	gcgccaccat	gccaggcttc	15420
ccatttgctt	tcaaccagac	aagtgaggcc	aggccaagag	cccaggagc	tggcgccctc	15480
gtacatttct	cccggcgtgc	acagggcacc	tcccaaacac	agcctgtgat	ggtgacacac	15540

gggctcccc	aggtcaagtg	gcaaagtctc	ccccagggaa	gaaaggagga	agccatgcct	15600
ggcaaaaagc	acacctctcc	tgcccaacgc	tttaacctct	gtatacaaat	caggccatgt	15660
gcactcgctc	cttcttacaa	tgctcataat	ttatactttc	agagtaaatg	aaacttggca	15720
tcaacccgag	aaacagctat	tcttttctag	atgcttacag	tgcccagcaa	atgaggactc	15780
gggtgtaatg	agattatgga	caactggaaac	aggatcataa	tgtgacgtgg	tcggtaaatgt	15840
gcagttttat	ttgcttaatg	accctcgccc	cgtgacaggc	tccctgaggg	tgggcctggg	15900
ggcagaggtc	cccgccacgt	ccccagccct	cagcacagtt	gccaggagag	ggtgacactc	15960
atgaagtggc	acagggaaga	tgggagctgt	gggctctgca	gatccaccac	ctcttctgtt	16020
catttttgtt	gatgctgttt	tttaagaaaa	ttattgaagt	aaaattcaca	ggacatacgt	16080
ttactttttt	tttttttttt	ggagatgggg	tctcactctg	tcaccaggt	tggagtgcag	16140
tgggtgtgatc	tcagctcact	gcaacctctg	cctcccaggt	tcaagcgatt	ctcccacctc	16200
cgctccaga	gtagctggga	ccacaggcgt	gcaccaccac	accagctaa	tttttggggg	16260
gtatcttttt	ggtagagaca	gggtttcgcc	atgttgccca	aggctggtct	tgaagccctg	16320
agctcaggcg	atccaccgcg	cttggcctct	caaagtgtg	ggattacagg	cataagccac	16380
tgcaccacgc	ctaaattttac	cactttaaag	tgaatagtgt	tacctagtgc	attcgcaagg	16440
cggtgcagcc	tccactttctg	tctagtcca	aagcacttcc	attgccccac	aggcaaaccc	16500
cacaccggcg	agcagtcacg	ccccagtcct	cgccccagc	ccgggcaaac	acttttgatg	16560
gacttaacta	cacacattct	caacatctca	tataaacgga	atcacaatat	acagcctctg	16620
atgtctgtct	tctttgactt	ggcaccatgt	tttcgaggtt	catccaggct	gtagcatgtc	16680
agtgttcat	cccgtttttag	gggtgaacca	tattccagtg	tgacagacaga	aaccaatctg	16740
tgcattccat	cacccactgg	gggacctttg	tgtcatttcc	accctcggct	gttgtgcaca	16800
gtgtgtctac	ggacattact	gtccattcac	attttgtgtg	aagacctgtt	ttcgattctt	16860
aagagtatac	agctaggagc	ggaattgtct	ggtcatacgt	aaatcaatgt	ttacgtctca	16920
aggaatcaac	aaactgtttt	ccacaatgtt	gtcttttttg	ttgtttttct	gagacagggt	16980
cttgctctgt	cacccaggct	ggagtgcggt	ggtgtgatca	tggctcactg	cagcctcaat	17040
ctcctaagct	caatccatcc	tcctgcctca	gcctcctgag	tagctgggaa	cacaggatatg	17100
taccaccatg	gccagctaata	tttctaattt	tatttttttt	tgttttttgtt	tttttgagac	17160
agagtctcgc	tctgtcgccc	aggctggagt	gcagtgggtc	catctcagct	cactgcaagc	17220
tctgcctccc	gggttcacac	cattctcctg	cctcagcctc	ccgagtggct	gggactatag	17280
tcaccggcca	ccacgcctgg	ctaatttttt	tgtattttta	gtagagatgg	ggtttcaccg	17340
tgttaccag	gatgggtctg	atctcctaac	ttcatgatcc	acctgccttg	gcctcccaaa	17400
gttctgggat	tacaggcgtg	agccaccacg	cccgacctta	cttttaattt	tttaatttta	17460
ttatttttatt	ttattttttt	tttttttgag	acagagtctc	gctctgtagc	ccaggctgga	17520
gtgcagtggc	gggatctcag	ctcactgcaa	gctccacctc	ccagggtcac	gccattctcc	17580
tgcctcagcc	tcccagatag	ctgggactac	agggtgccac	cacgatgccc	ggctaatttt	17640
ttgtattttt	agtagagaca	gggtttcact	gtgttagcca	ggatgatctc	aatctcctga	17700
cctcgtgatc	cgcccgctctc	agcctcccaa	agtgtgggga	ttacaggcgt	gagccaccgc	17760
gcccagcctt	tttttttttt	tttttttttt	ttttgagata	gagtcttgct	ctgtcgccca	17820
ggctggagtg	cagtggcggg	atctcagctc	actgcaagct	ccgcctccca	ggttcacgcc	17880
attctcctgc	ctcagcctcc	cgagttagctg	ggactacagg	cacccaccac	cacacctggc	17940
taatgttttg	tattttttagt	agagacgagg	tttcaccgtg	ttagccagga	tgggtctcgat	18000
ctcctgacct	cgtaatccgc	ccgcctcggc	ctcccaaagt	gctgggatta	cacgcgtaag	18060
ccatggcgcc	cagcccattgt	ggccattttt	cagtgaagga	agccagaggc	ccatcactct	18120
cggttgctcc	ctgggcatag	ctctgcctca	gccagaagca	ctgagggaag	gtcagcctcg	18180
gcccttgccc	cagccacagt	cacagataaa	ggggcctgca	caggctctgtg	tggctccaga	18240
gctcgtcacc	caacacacga	cgcttccatg	tgaatagccc	cagggtgcatc	atgaagagcg	18300
atggccgctg	cagaggcaga	agaatcccg	ggggaagcag	gtgggagaga	ggctgagaac	18360
agaccagacc	ctggagctac	agaccctatg	ttccaaccct	ggctgggact	agctgtgtgg	18420
ctctgggcaa	attcacatgc	ttctctgtgc	acaggggatc	aaaatagcaa	acacaggcta	18480
ggcacagtgg	ttcacacct	taatcccagt	gctttgagag	gccgaggtgg	acacatggct	18540
taagctcagg	agtttgagac	cagcctgggc	aacatggtga	aacctcgtct	ctacaaaaaa	18600
aataccaaat	aaattagcca	ggcgtgggtg	tacgtgcctg	tggctctcagc	tacttggaag	18660
gctgaggcgg	gaggaacact	tgagcccaag	aagtcaaggc	tgtggccgcg	tgtggtggct	18720
cacgcctgta	atcccagcac	tttgagaggc	tcagggtgggt	ggatcacttg	tgatcaggag	18780

ttcaagacca	gcctggccaa	catggtgaaa	ccccgtccct	actaaaaaaaa	tacaacaatt	18840
tgccaggcgt	ggtggcgggc	acctgtaatc	ccagctactt	gggagggtga	ggcaggagaa	18900
tagttagaac	ttgggagggtg	gaggtttag	ttagccaaga	tggtgccgct	gcactccagc	18960
cagggggaca	gagcaagact	ccatcccaaa	aaaaaaaaaa	acaaacaaac	aaacaaaaaa	19020
agaggtcaag	gctgcagtga	accatgattg	tgccaatgca	ctccagcctg	ggtgacaaag	19080
tgagaccctg	cctcaaaaaca	ataaaaaatat	aaataaaaaat	aaaacataat	agcaaactgt	19140
tcatagaggt	ggtatgagca	ttaaatgaac	tgataaacgt	ccctggaaaa	cagtaagtgc	19200
tatggaagga	ttcgtgccc	ccaccgccac	caccattagc	atgtttcaac	ctccatcacc	19260
ctcactgtcc	cctgtcacca	tcctttgacc	agggcactcc	cagctgcagc	ctttctatcc	19320
tcttgtccac	ccttcataac	tgtaagatca	ctcagctccc	aagaaccaca	gtctacaggg	19380
taaccacatt	tccaaatctc	aaaccagacc	cgctgggtctg	cacttccagg	gacaacagga	19440
tattttcaaa	ccagcccaaa	agagatgtgt	ggctcagcat	aagaggaaca	ggagaaaactg	19500
aggcctcttg	ccctgagaat	gagcttgga	gtggatgtcc	cggcctcact	caaaccttca	19560
gatgactgag	gcccagccag	gagcttgagt	gtaccctcag	gtcataccct	gagccagaag	19620
caccagctga	atccactcct	catcactgac	tccctcccca	taaaaaacct	gtttgctgtt	19680
tcaggctgtt	aagttgtggg	ctgttttgtt	acacagcaat	ggataactaa	cacacgagggc	19740
ctggcaagt	tgagcaaaag	ctgcccagc	cctcaagtct	gttcatgtgg	gtgttggcct	19800
gtgtttgcag	aaatccagcc	actgagtcct	cccatgcagt	cactactgcc	ctctgcacag	19860
acacctgcca	catccctgcc	tgggcccagga	gctccactag	tgcaggaatg	gggtctgccg	19920
tcccaggagg	atccctgaca	cctagcacag	ggctagcagc	aggcagcact	tggttagtga	19980
ataaactgcc	cttcacctgt	acacagaagg	gatgtttcta	taaggggtaa	ttaagtacag	20040
agctgggaag	ctatgctgac	cagaaggctc	taaaagcaat	taaccaacga	ggggaaaacc	20100
cttcctactc	attctcggcc	cattttattg	agcactgacc	atgtggaagg	ccccctggtg	20160
agactgggga	atgcaccaat	aactgagaca	gcttccggct	gttgccctca	ggatgcctga	20220
gctgggatag	ggccagggtg	gggggtggtc	gtgtgacagg	gttactgttc	acaacctctgc	20280
cgggcataaa	gccctcccca	acaattccaa	aatccaaaac	gctctgaaga	tggaagctt	20340
ttgttgctca	tctggtgaca	aaacctcatt	tggtgcatgg	gccgggtgcg	gtggctcacg	20400
cctgtaatcc	cagcactctg	ggagccgagg	ggaaggatcc	cttgagctta	ggagtttgag	20460
accagcctga	gcaacatgtg	agaccccgctc	tctacaaaaa	atacaaaaat	tagccagggtg	20520
tggtggcgca	ctcctgtagt	cccagctact	cgggaggctg	aggcgggagg	atcgcttgag	20580
cctgggaggt	gggggctgca	gtgagctgag	attatgacat	tgcactccag	cctgggtgaa	20640
agagtggagac	tctgtctcaa	aaaaacaaag	ttaaaaaaaa	aaaaactgtg	catgggtgtg	20700
ggctacagat	agtcttttct	gccctactta	gaatgaacgt	gccacatttg	ctatagaaat	20760
attcaagggc	tggtggcaaa	tgccacacag	accctgacgc	tgttccaagt	tctgagaagt	20820
cctgcattcc	tcaggggccc	agagtttcag	agaagagtct	gtaggcctga	gttaagaagg	20880
aacgccttca	aaagccctgg	ggacaaaggg	gaaaggggtg	ccccaggact	gcgtgggtac	20940
ctaccggaac	gagccgtcca	ggttggcacg	gtggatgaag	ctgagcttgg	cgtcagccca	21000
gtagagcttc	tgtcctcca	ggtcgatggt	cagtccattg	ggccagtaaa	tgtccgagtc	21060
cacaatgatc	ttccgggtgc	tgccatccat	ccctgcccgc	tcaatccggg	gcgtctcacc	21120
ccagtctgtc	cagtacatgt	acctgtgacg	ggggcagggc	aagagaagca	gctaacacag	21180
atctgttttt	tgtttttgtc	tgcatagatg	cagacatgaa	acaacagaca	gtgaacttgc	21240
cctaaaatct	cacccatcgg	aaataaccaa	caggtatggt	ttcagggtatt	cctgccttaa	21300
gctgggcaat	caaaaatatac	tatttccaac	ttgttctcag	ttaacagtaa	attctgggca	21360
ccttcccttc	ttgtggatag	aaagattcct	tggttctttg	atgattgcct	agtgtactct	21420
gctgtaagtt	ttttaaaagaa	cttcagggtta	tttctgattt	ttttgctacc	atgaaaatgc	21480
tgtaaatgaa	cctctaaaag	gcaattcaaa	acactcagga	tggaaatatta	tttagtggtta	21540
taaagaaatg	agctatcggc	tgggcccagt	ggctcacacc	tctaatacca	gcactttggg	21600
aggccaaggc	gggtggatca	cgaggtcggg	agatcaagac	catcctggct	aacacagtga	21660
aaccccgctc	ctactaaaaa	tacaaaacat	tagccaggcg	tggttagtgag	cacctgtagt	21720
cccagctact	taggaggctg	aggcaggaga	atcatttgaa	cccgggaggg	ggagggtgca	21780
gtgagcagaa	atcgcacccat	tgcactccat	cctgggcgac	agagcgagac	tccatctcaa	21840
aaaaaaaaaa	aagaaaagaa	aagaaatgat	ctatcaagcc	atgaaaagac	atggaggaaa	21900
cttaaatgca	tggttagtagg	tgaagagacc	aatctgtatg	agtccagttc	taaacactct	21960
ggaaaaagca	aatacacaga	gacagtaaag	catcagtggt	tgccaggagt	tgagaggagg	22020

agggatgaat	gagtggagca	cagaaaatca	gggcagtgga	actatcctgt	atgacatgga	22080
atgggtgggtg	catgtcctta	ctcatctgtc	taaaccaaga	atgtacaaat	caagggcgaa	22140
ccctcgtgta	aacgtggatt	ttgggtgatg	gtgcgtcagc	cagctttcat	cagttgtaac	22200
aaatgtacca	ccctgcacag	gatgctgaca	gttgggaagg	ctgtgtgggt	gtgaggacag	22260
ggatgtatag	gaactcagta	cctgctgctc	atcaattttg	ctgtgaacct	acaactgttt	22320
gaaaaaatta	agtctattta	aaaacaacaa	aacatggcca	ggcacgatgg	cttgccacctg	22380
taattccagt	acttcgggag	gctgaggtgg	gtgggtcact	tgagccaccc	tgggcaacat	22440
ggcaaaatcc	cacctctaca	aaaaataaaa	attaaaaaaa	agttagctgg	gcatgggtggc	22500
acactcctgt	agtcccagct	acttgggagg	ctgacgtggg	aggatccctt	cagccctggg	22560
aggtcgaggc	tgcagtgagc	tgtgactgta	ccactgcact	ccagcctgga	tgacagagtg	22620
agaccctgcc	taaaaaaaaa	aaaaaaaaagg	ctgggtgctg	tggctcatgc	ctgtaattcc	22680
agcgcttttg	gaggccgaga	tgggccggtc	acgaggtcag	gagatcgaga	ccatcctggc	22740
taacacgggtg	aaaccccgct	tctactaaaa	gtacaaaaaa	aaaaattagc	cgggcatggt	22800
ggcggacacc	tgtagtcaaa	gctactcggg	aggctgaggc	aggagaatgg	cgtgaaccgc	22860
ggaggcggag	cttgacgtga	gccaagatca	caccactgca	ctctcagcct	gggagacagc	22920
aacactccgt	ctcaaaaaaa	aaagaataaa	acccatgggt	gggatggacc	ctgaacctgc	22980
agctgcagct	gttctctggg	aggctctgtg	gcgacgtggc	tttgcttctc	catgttccca	23040
agagacaagc	atcacccatc	catgagaaac	agcacatccc	tcaggggcgc	cttacgtgat	23100
ctctggccaa	tgaaccaaga	caaagtgaac	agacaccagg	tctgggatgg	caggtcccac	23160
ccccaccagt	gcccagtgtg	ccctgttttg	agggtgaccac	agggtgtgtg	cccagaggct	23220
gggcgtgact	ctcagcggag	accagagggg	aaccacacca	gcttggagga	ctcagttccc	23280
atcccagcca	gctgggatga	gccacaggac	acaagggtcg	gcagacctat	tgtgttttgt	23340
ccacccttca	cagcagagaa	aggggacagt	gcccagaatg	tcctctgagg	agcctcctcc	23400
cactcttgggt	ccttgtaaaa	tgggtgctgac	tcccttgctc	ccttcttccc	gggggtgggcg	23460
gcaaacccca	ttccctcag	ccttagcaag	tgatttagaa	acaggcagct	cgcccaagcc	23520
aggcatgaga	gtgatcccgc	gacacaggga	gaacaagccc	cgctttgccc	tctgggggtc	23580
tccattcagc	agaagaggca	aatgacagac	acacagccgc	ctcctcccc	accatgggtgc	23640
tctgcagcct	caggagcctc	agggtgcacca	agggccaccc	catccagggg	gccatgcttc	23700
cttgagtgggt	atcgttcctg	agcgagtacc	atctccacct	tccagagggg	ctgtgacaag	23760
atcaacaaga	atgagggcat	aggagcctcg	aaccaaaccat	gccctcttcc	ctgcagaggc	23820
tgactgcgcc	cagctgctat	caccaagccc	ctgctcctcc	ggccccgtgg	ggacagggtg	23880
agaggggtgt	cacatggaac	agctctccaa	acagtccctc	tcaagctgct	gtctcctgtg	23940
catctagtga	gaacccaacc	aacaaaggga	agggtgggaat	tgctattccc	attaggcaga	24000
tgagaaaact	gaggccccga	aaggctggcc	tgttccaggt	tacaggcgct	gagcggctgc	24060
tctgggaaca	cacttgggtg	ctgctgaggg	cccgagcccg	gccatcata	gactcaccct	24120
tgcgcagcaa	agcccgggtg	tgggtgaact	tttctgggca	gcctgggact	ccaaggtgct	24180
ggcagccagc	ccagggaagg	ctcccgcgtg	cctgcgccag	acgccttgct	ttacctgcac	24240
gtccccaccc	ctaggagcct	ggacagagcc	cagaccctcc	gccacctcct	gagaaggat	24300
caggggcatc	agtctggact	tgggggggaa	tccacacagg	ccttcccca	atgctccacc	24360
gtggcccatg	gaaaaggctg	gaaaacgtgc	aggagcagga	gcctccgcat	ggagcataat	24420
tcacattcct	tccccgagtt	tcataacaga	ggcctgctgg	tttccctaaa	tggggaattt	24480
gcgagccagt	cggtgaccag	agactgggtg	gcgtggacgt	gctcttgcat	agtctcaaac	24540
gctaccacaa	gcccagccaa	attccacgga	ggaaaatcga	cttccgaaga	aaagagctgc	24600
agcatggcct	togtgcagag	ccagctgcgg	ttgtgggtgt	gtgttatatt	aggggaagggc	24660
cattttgcat	tttaaagagg	gggttgggtt	tcacctggc	tttaatttga	gacccggggg	24720
ccactgcagc	cccttgtcag	gctggtacag	gccggggact	cctcccatgc	taagccagtgc	24780
tctttctggc	cccagatcct	caggggcccag	agggtcaccc	ccagagcccg	ctctgccacc	24840
cacatgggtg	ccctgggcct	gggaggggatg	tgccttccct	caacctgcc	tggatgtccg	24900
cacggggcca	cctgcattgc	tgaaactgca	acgaagtoga	gtctcaggag	gggccccctt	24960
ggctgcaggg	ctcttgatcc	ttttggccac	gtgcacactg	agggtggacgc	tgggacccag	25020
agacccccctt	catgatgatg	gcccggggcag	gaacccccctc	ctctgaggaa	ggacccctggt	25080
gggggacagc	actgcaggag	ggcacaggag	atgacggggg	ctctagcagg	gcccgggagga	25140
aggccaagat	gctcctcgca	accgtgtgcc	tgtggccagg	acagaggaca	aacccacctt	25200
ccactgtccc	cactctcagg	acagcagttc	tgccccagga	ctcagcgccc	acacttatgc	25260

ctgaggacca	ctattcaagt	cagtatttgg	cgagcagggg	ttgctgccgc	gggcgctgtg	25320
acaggctgga	atcctctccc	tctccctctc	cctctccgga	gacatggagc	ctacagggac	25380
agagtcagca	cctcagggtg	ggaccatggc	tggcgtcatc	agcatcactg	gatctgatga	25440
gtgggagccg	gcatctcact	gttttcactc	tctcattcaa	atgactggag	caaagggag	25500
gtgtggggag	aggcccagga	atcaacacta	aggtcaactt	tgccccaggg	ggcaggggtg	25560
ggagtgaaca	gccacaggtg	tgatcctggg	gagggcttct	gggagagaat	tcagaggcaa	25620
gcatgtagag	gaaccatttc	aaatagttaa	gaaaagccag	agccaaacag	ggacagttgg	25680
ctcgcagaga	tgatgcaggc	aaagccagct	cagatctgag	catgggaaaag	actactccca	25740
accaagggcc	cagcatctcc	caaccaagca	ccaagtacct	cccaaccaaa	tgccaagcac	25800
ctcccaatca	aatacctccc	aaccaagcac	ctagcacctc	tcaactggac	accaactact	25860
cccaaccagg	caccaagtac	ctcccaacca	agtgccaaagc	acctcccaac	caagtaccaa	25920
ttacctccca	accaagcgcc	tagcacctcc	caactgagca	tcattgcacct	cccaacagag	25980
catctagcac	ctcccaactg	atcacctccc	aacctagcac	cgagcacctc	ccaaccaagt	26040
gcagagcacc	tcccaaccac	gtgccaaagc	cctcccaatc	aaatacctcc	caaccaagca	26100
cctagcacct	ctcaactgga	caccaacaac	tcccaaccac	gcgccaagca	cctcctaaca	26160
aagtaccaat	caccttccaa	ccgagcacct	agcacctccc	aactgagcat	catgcacctc	26220
ccaacaaatc	acctagcacc	tcccgactga	tcacctccca	acctagcact	gagcacttcc	26280
caaccaacat	agcaaaagcc	ataaagaagt	aaaaagacaa	aaccacgtag	gcatggagac	26340
tggacttctg	gtggcgagga	aagggcattt	ttattataac	gacagctaac	atttggtgaa	26400
ctcacaaact	gttcttggtg	tttctctcat	gacatgcagc	atgggtcacgc	ctctgtacag	26460
acaaggatac	tgaggcagag	agtggcaccg	tgccaacctt	gtctcatctt	tttatcgaa	26520
ctacatgcag	agtgccagca	aatccagctg	tcttttctct	tcagaacaga	tcccaaatct	26580
cgccactcct	tacccccaca	agtgaggtgt	ccccgctget	gctttctgtc	gccaggatcc	26640
cggtataaac	cgtggagagg	gtcctgccc	ccacgccacc	cacccacag	ctcactctcg	26700
ctccagccac	caggggatgc	cttcagcac	gagtcagagc	tggcacctcc	tctgctcgag	26760
acctcatgtg	tcctctcctc	acacctggg	ccctgtttcc	ctacattctg	ctacagcccc	26820
tcaaacaggc	cccgccccaa	accagcccag	ggcctttgca	ctggctgatc	cctctgcctg	26880
gaccgcgctg	ccccagaca	gccacacggt	tctcagcctc	atctgcttcc	agtctcgact	26940
caaaagtcac	caagaggcct	tcccagcacc	tgagctccga	cggaagcccc	tcgccacagc	27000
accaagcac	tgctttatcc	ccctacgcac	acgtcccttt	caaatactat	tcatattacca	27060
tctcctccca	ctcactgaaa	gggccagaga	ctgggctata	cccgctgctg	ggggagcagg	27120
accaggcgca	agggtcacca	aatgcagtg	atgcctgggt	gggaggtgag	ggagctgcag	27180
cgaccacgc	tgggagggaa	cgcaatgaca	ggaggagcgc	aggtcctggc	gacacgatgg	27240
ccatggcagc	cgctggtgag	caaccgcagg	ccggccctgg	gagagggctt	ctagcaagct	27300
gctatcttca	gcctctccga	ctactgcaga	tgccccctcc	tagccagaga	cactgctaca	27360
ccagccgacc	cttccaaaaa	gaaggtcagt	aacccgcga	ctcctggagc	cacagtgcag	27420
ggggagaggg	ctgagagggc	aacagttcac	caagcggaac	agaggctgcc	ccggaggtca	27480
gctggctccc	cggcagctgc	aggggtggct	agcccactcg	gagggcagcg	agggcatacg	27540
aggggtcca	gggatgagt	gttgcccagc	acagcacccc	tgggagggccg	ggggcacttc	27600
tcaggtagtg	ggggcagag	gctgctctgg	cctgacctca	gggactcaaa	atactttggc	27660
gataaattcc	accgtgtccc	accctgctg	gtaccccata	cttacacaca	gactgggttca	27720
gatgcagaca	ctctcgcgca	catactcgct	cacacgggca	catacacgtg	cacacacagt	27780
cacatgcgca	cactcataca	cacacaaata	tccactcaca	cgcattgcag	cacacacag	27840
gacacacaca	ggctcacacg	tatgcacgca	tatgcgtgca	cacgcacaca	cacacacaca	27900
cgctcacatc	ctcccactcc	cacactcagt	tgtcagaca	cacacacgcc	tggctctcac	27960
acaaacctgt	tgggctctga	aaggctccag	cccttcccat	gctcgtcaga	agccagtcaa	28020
tggcttccca	agtcaccaca	cagatcaaag	aggtgaactt	ggccacatgg	cactctgctt	28080
cctgagctcc	caaacaccag	ccttggtgag	gacagaccct	cacccacac	cctcattccc	28140
actaccctgg	gcaggcccag	aggaggggca	tctgcaggat	ctggcaacca	gcccctcccg	28200
cccggtcct	gcagccggca	ccatgggagt	cagggggagg	tcactgcaaa	gggcaacagc	28260
aagtgggtgg	ccccaggact	agagcccagg	ggctctcagt	cctactccag	agcttggaca	28320
ctgtcccaca	gggcatggcc	aagggaaagg	cttccagagc	cctgacttca	gggagggagg	28380
caggcgggct	cctgtggcag	gcctggatgc	atggccgccc	actcctggga	ctttctaacc	28440
tagaatatct	aggtcaggct	gggtgcagtg	gctcacgcct	gcaatcccaa	cactttggga	28500

ggccgaggag	ggtggatcac	ttgaggttag	gagtttgaga	ccagcctggc	caacatggcg	28560
aaaccctgtg	tctactaaaa	atacaaaacc	tagccagggtg	tggtagtgca	cgctgtaat	28620
cacagctact	caggaggctg	aggcaggaga	atcacttgaa	ctcgggaggt	ggaggttgca	28680
gtgagctgag	atcgtgccat	tgcgcaaaga	agatctaggc	cggccctca	accggtgagg	28740
tccaggctgg	gagtgtgag	agactgtggt	gacactgaat	gaactaacag	gcaaagggct	28800
tccaactgag	cctgggggtg	gtgggaaatg	gctcttgtgt	tctagtcaag	acctctgcca	28860
accagttctg	acactgaccc	agcacagaac	ctgacagggtc	agcaagggcc	agggcttagc	28920
acagcccagg	taagggtgtg	tgtacggccc	ccagagtcac	tcccaggctg	caagaaaagg	28980
gacaaaggag	ggacaagggg	tggccaagca	aactgttccc	tctgctcggg	agtctggggg	29040
gacctggcct	agctggccag	tggagctggg	ccacctcccc	ttaaactctc	caccccgga	29100
ttcgactcca	aagctttcct	gccacccacg	ctctccccac	ctgggatcac	ggccaggccc	29160
tgagccttca	aggcccagg	tgaactcagc	cagactagga	gctgaggagg	acacagggca	29220
gcttccagaa	cggacccgag	aaccactccc	agcaggttct	gcttccagac	aaggagctgc	29280
actttttcag	ccaatgcaat	tagaaagcca	ggagaagggtg	caaattccac	ctgctgagc	29340
gtccgcactt	cccaggccgc	ccaccataca	cacagcaaag	atgtgtttta	ccattcaaac	29400
ccatggccaa	ccacatcggg	tgcctcagac	atgcaagttt	taaaaaggaa	cataactatg	29460
ggccaggcac	ggtggttcac	gtctgtaatc	ccagcacttt	gggaggccga	ggtgggtgga	29520
tcacctgagg	tcaggagttc	gagaccagcc	tagacaccat	ggtgaaaccc	catctgtacc	29580
aaaactacaa	aaattagctg	ggcgtggtgg	tgggcgcctg	taatcccagc	tacttgggaa	29640
gctgaggcag	gagaatcact	tgaacccggg	aggcgaagggt	tgcagtgagc	cgagattgtg	29700
ccactgcact	ccagcctggg	caacaaggga	gactccatct	caattaaaaa	aaaaaaaaaa	29760
aaaaaggaac	ataactatgg	agtctcaagg	ggaagtaatt	ccttcaacaa	taacaaatct	29820
tgaaagctga	gctctttttt	ttttttgaga	caggatctcc	tcactttgtc	gcccaggctg	29880
gagtgcagtg	gtgggatcac	agctcactgc	agcctcgatc	tcccaggctc	aatgatcct	29940
cctacctcag	cctcccaaga	agctgggatt	acagggtgat	accatcacac	ccgattcatt	30000
tttgatact	ttgaagagat	ggggctctac	catgttgccc	agtgtggtct	tgaattcctg	30060
gactcagggtg	atctgcccgc	cttggcctcc	cagagtgtctg	ggattacagg	cctgagccaa	30120
cacccccacg	ggttcatttt	cagagtgcga	ccgagtgtctg	gggttacagg	cctgagccaa	30180
ccccccccacg	ggttcatttt	aagagtgcga	ccgagtgtctg	gggttacagg	cctgaaccaa	30240
ccccccccacg	agttcatttt	cagagtgcga	ccgagtgtctg	gggttacagg	cctgagccaa	30300
ccccccccacg	ggttcatttt	aagagtgcga	ccgagtgtctg	gggttacagg	cctgagccaa	30360
cacccccacg	ggttcatttt	cagagtgcga	ccgagtgtctg	gggttacagg	cctgagccaa	30420
ccccccccacg	ggttcatttt	cagagtgcga	ccctttttct	gaaaaacaac	ttgggctcat	30480
gcaaattcga	gagagagatg	gtgacactcc	ccgccccctg	gacccagggtg	gagtcgcagc	30540
agggtttacc	cgtgagcggg	gtccaaggcg	atggccctcg	gctggtcaag	gtcctgccag	30600
aagagcacct	tccgggatgt	gccattgagg	ttggccacct	cgatgcgggt	ggtctctgag	30660
tccgtccagt	acaggtctct	gccacccag	tgcagggcga	ggccgtcggg	agagaccagg	30720
ccggagatga	ccacgttctg	cacggcggcc	cccgtctggt	tcaggtaggt	ctgcttgatg	30780
gcctcctcgc	tcacgtctgt	ccagtacacg	gctcccttgg	aaaactggaa	gtccactgcg	30840
gccgcacctc	ccaggccgct	gaccacgatg	gtggactcca	gcttgactcc	gccggcgctc	30900
accagccgta	cgtcccggcg	gttggcaaat	agcaggagcg	gcgaggctgt	ggggcagaag	30960
caaaccgtga	ggggcactgg	ctaagccagc	aagatacaca	gccctgggat	ggagcactat	31020
gcccagagca	ctcctggtac	tgcctgccc	atgcccaaga	cctccagttc	cttctctcca	31080
cccctaaggc	gttgtcagga	agttgcctgg	gcagccccgg	cccgcacatc	tcagaggctc	31140
ctgcagcgca	gcaaacagcc	ttcttcccac	attcggtgac	agcacctgtt	tgtttaccaa	31200
ctgttacgtc	tgttccccca	gatatgggtg	acccttctctg	ccatgcccaa	aacctcccac	31260
atcgtcctcc	agaggctaca	ggggccctgt	cctgttctgc	agagaagcca	catccccctt	31320
gttggcctga	cacaggggat	ggggacatgc	aggcacagca	ctggccatgc	tgctcgctac	31380
agaccagcc	acagggccac	atgtttttgag	gggttcagag	cccaggccag	acagagcctc	31440
aagattccct	tacaagtctt	tgaccactgt	ccaagctcag	gcccgtttcc	ttggccgtgg	31500
catcagcttc	ccatccaccc	ctgtattcca	tgtttctccc	accctgcttc	tggacattcc	31560
tacattttaa	gggtcactct	ggaatgccac	cccttggttc	agacaccttc	cacagctccc	31620
tgtgccagtg	ccatgcagaa	caaggtcaga	ccccctagcc	tggcctccaa	ggccttgccc	31680
tctggcctca	cctacacttc	tctccaccac	cccaccccaa	gcattcctga	tctgcctgcg	31740

gccaggctgg	ctccctcacc	tccctgtgca	ccgcagccct	cagccccttc	tgccctgtgca	31800
agaagcctca	tctcacagac	aacgggtctca	ttcccacaac	gggctcaatg	agaaatcagg	31860
agaggccttc	agaccatcac	cccaccagac	acctcagacg	tcggaccagg	aggggtccagc	31920
aacccccaac	acagactcag	agggactaag	aagccacatg	aggagtgaac	acaagatgtg	31980
gacaggagga	gggttaagggc	ctccaggggag	ctccatcagt	ccgtgttctg	ctgtcagcag	32040
gggttaggctg	ggctggccac	aaacaccccc	aaaaaacatc	tgaagccttg	gcttgaaaca	32100
gctgacattc	ctcatgaaaa	ctgcagaccc	ctgggtcctc	ctgcgcagat	gggggagccc	32160
agccaacccc	acactcccac	cttcaccaag	aaagagaaaag	ccaaaacaaa	ctcaactcag	32220
ccaatgacaa	tcacagaact	gaatcctgta	gttagttcag	ttggtttcat	ttcagcaggg	32280
gaaagatttg	cagcctctat	gagggtagct	gggaacacaa	agggccagag	catggcccag	32340
gagaccccag	cgcagtgggg	tagatgggtc	cgagcacagg	cctccctgcc	aagacaagca	32400
ctggctcaaa	tcctggcccc	tcccattccc	aggagacatg	ctccacagga	tgaggaggaca	32460
cacagaggac	ctgaggccag	gaaaatgaca	gcggcgccctc	cgccgccccca	cccgtgctgt	32520
catcatctta	gggtctacagt	tctttgtggc	aacgagggac	actgtgaaag	tcaaacaaca	32580
ggaaggcata	ggccacaaat	aaagacaaac	gggacttcat	gggaagctaa	agattttgtg	32640
catcaaaaga	cactatcgag	agagtaaaaa	ggcaaccacac	agaatgagag	aaaatatttc	32700
caaatcatag	atctactaag	agattaatat	ccatgaaata	cagagaactc	ctaaaactca	32760
acaatgagaa	aacaactaag	ccaactcaaa	aatgggcaaa	caacttgaac	agacatttct	32820
ccaaagatga	catataaatg	gccaataaac	acatcaaaac	aggcttaata	tatccctaata	32880
catcagggaa	atgcaaatca	aaactacaat	aagataccat	cttgaccaa	ttaggacggc	32940
tactatcaaa	aaaacaaaat	agcaagtgtt	ggtgaggatc	tggagcaact	ggaacccttg	33000
tgaccacttg	gcaaaaatgt	gaaatgggtg	agctactatg	gaaaacagca	tggcagttcc	33060
ccaaaaactt	aaacacagaa	ttaccatatg	acccagcaat	ttcgcttttg	gttatataacc	33120
caaaagaact	gaaaacaggg	acacaatcag	atatgcatac	accttggatc	acagcagcat	33180
ccttcccaac	agctaaaaca	tggaggcgagc	caggcatggg	ggctcacgcc	tgtaatccca	33240
gcactttggg	aggctgaggc	gggtggatca	cctgaggtca	ggagtccgag	accagcctgg	33300
ccaacatggg	gaaaccccg	ctctactaaa	atacaaaaat	tagctgggag	tagtgacggg	33360
cacctgtaat	cccagctact	cacaagtctg	aggcaggaga	atcacttgaa	ccctggaagt	33420
ggacgttgca	gtgagccaag	attgcgccac	tgcattccag	cctgggtgac	acagcgagac	33480
tctgtctcaa	aaaacagcaa	aacaaaaaca	aaaaaacaaa	caaacatgga	agcaacccaa	33540
gcgtccctct	actgagggat	gaatagcggg	gcaaaatctg	ctccatccac	acaatggagt	33600
actattcagt	ctcaaaaagg	aaaaagattc	tggtcaggca	cgggtggctca	tgccgtgtaat	33660
cccagcactt	ggggaggctg	aggcgggtgg	atcacctgaa	gtcaggaatt	caaggcccg	33720
ctggccaaga	ctggcaccna	gctacacana	aagtatangg	ccccggaaa		33769

<210> 9

<211> 72049

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (8356), (8385), (38585)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 9

tataccttgc	gcggaccttc	ggctcctgtg	gtgaagacaa	tatgaagaaa	atagaaatta	60
cccataatit	tgccacacag	acttagttgt	gtccatgtat	cttgtgcacc	ttttttctgt	120
ttacggatca	aaatcgactt	ttagggtcag	gcgcgggtggc	tcacacctgt	aatcccaaca	180
ctttgggagg	ctggagttgg	ggttgggggg	tggatcactg	aagatcagga	gtttgagacc	240
agcctggcca	acatggcgaa	actccatctc	tactaaaaat	aaaagattag	ccaggcgtgg	300
tgggtgggtg	ctctaattccc	agctactccg	gaggtcgagg	caggagaatc	gcttgaaccc	360
aggagacaga	ggttgcagtg	agccaggatc	acgccactgc	actccagcct	ggcaacagag	420
cgagactctg	tctcaaaaaa	aaaaataaaa	ataaaaataaa	taaatacata	aattgacttt	480

taggagattg	gttcaaacaa	tgtgtgtaat	gttgtgtctg	agtgtttttc	atttatcggt	540
catgcaaatt	ccgacatcat	tcactcttct	ccagagtgtg	ctgttttcc	gcctgtgtca	600
tcacccgtca	ccttgaatgc	cctcgtttag	gtaaaataag	tacattttat	tcaaaaatat	660
ttgaggacat	ttgggttgct	tccaggttct	tggctctgag	ttttgctgtt	cttgtggagc	720
catggtgggt	tctgggtgca	ggaacctcca	tgcgttccag	ctgctgcttc	tgcctgtgtt	780
cttagagagg	aaatgctggg	gtccgcggtt	cccgggctgc	tgaccaggaa	gcctgcgggtg	840
ctttacggcc	cttccagaag	cgggagatgc	ccccacttaa	gtgtcagaca	ggcctttcca	900
cctcactggc	agctctgagc	ggctcccttc	tatttgcaga	tgactgagaa	gttaccaatt	960
tccacgttta	ctgactgctg	tttctcctgt	taatttgtat	ttatagtctt	cgctaattta	1020
ttgctaggg	tttgggtgtt	tccctattga	cttgtatgcc	ttttaatttt	ttaaacaaca	1080
ttaataact	tcattttttt	agagcagttt	taagtttaca	ggaaaattaa	gggacaagta	1140
cagagagttc	cttccacctg	ctgtcctcct	ctcctcctcc	ccaccttccc	tccttccccct	1200
attgtaactt	tctttctgat	attataaaa	tcactcatgg	ctgggctgtg	tggctcacgc	1260
ctgtaatccc	agcacgttgg	gaggcagagg	caggcagatc	acctgagggt	aggagttcca	1320
gaccagcctg	gccaacatgg	tgaaccccg	tctctactaa	aaacacaaaa	agtttagccag	1380
gcgtgggtgg	gggcacctgt	aatcccagct	actcaggagg	ctgaggcagg	agaatggcgt	1440
gaacctggga	ggcagagggt	acagtgcgtc	gagatcgctc	cactgcactc	cagcctgggc	1500
aataagagt	aagcttcgtc	tcaaaaacaa	agtccacac	gcttcttgta	cgaggggtcat	1560
ttggccgagg	ggccagatgg	ctcaccatct	agttgggaca	ggccatgagc	tcggaatgct	1620
ttttacatat	ttacatgggt	gagaagaaaa	tcaggagaat	aatgttttgg	gacatgggaa	1680
aatgacatgg	aatttgcatt	ttagtgtcca	taaatgaagt	tttgtttgct	cccagctgtg	1740
ttgactgagg	caggctggct	tcctacagct	gcccgcagagc	tgaggaggcg	ggaaggagac	1800
cgtgcaggcc	gcagcaccga	aaatatttgc	tctctggccc	ttcccagagt	gcttgccgac	1860
ctctgtccga	cagctagaag	gaaggatagg	acccgtccga	cgataaccac	tgttgacatt	1920
tgagcgcgtt	tccttcccgg	cttttgtgtg	agagtggcag	tctgtttgct	tttgtggctg	1980
ggatctgctg	cacgcacggc	gggctgtttg	catgaggctt	cctggaggat	agggctgggc	2040
tcggagctgc	acgcagtggg	gcgtgtcctg	catgcagtgg	ggcctcagaa	gagagctgtg	2100
gtgggcgggg	cagtgcacac	gctgggtggg	gccaggcctc	cacgctcaga	tcagccccgg	2160
cgacagggtt	gggccaccct	ctctctggcc	tctgtgcagt	ggcccaggcc	gtctgctctg	2220
cctggcacac	ttgcctctgt	ccttccactg	aagcgctcct	cttaccctct	gctcccggct	2280
gggtacgttg	aattgtgtcc	ctcaaggaga	tatgctaaag	gtctaaccct	aggaaacctgt	2340
gtatgtgatc	taatttggaa	acagggtcct	ggctgatgta	atcaagcgag	gatgagggtca	2400
ccctagagta	gggggcctat	atccacgggt	ctgggtgtcct	catgagagca	ggtgagcaga	2460
cactgacact	caggggtgaa	ggctgcattg	agtcagaaca	gggcttagtg	cgatggcggc	2520
cacaagccaa	ggaactccaa	gtatttcctg	caacaccaga	agctggaaga	tgccagggaag	2580
gatcctgccc	tggagccttc	ggagggagtc	tgtccctgca	gacgtcttga	cttttgattg	2640
cagggatgca	tgtcttaggg	tgtgtggggg	ggtgcatttc	tgatgttaga	agccacctgg	2700
ttggtggcga	tgtgtcacgg	gagccctctg	cagggttctgc	gtgtccatgt	ggtcggggac	2760
agaggtgggc	agggacggac	ggtgtcgagc	tggacatgtc	catgacgtcg	gccatccctt	2820
gggatggcct	ttttgttttg	aggataaggc	tgcctgccag	gaagctgtgc	cctgcctggc	2880
ccttgcccca	agccctggc	ctgtgcttgg	cctcgcggaa	gggatgtcgc	ccttctctcc	2940
tgcattgcgt	cagggaggaa	ggggagaggt	cagcagcccg	cctggaggag	gctcgggcga	3000
ggggaagggt	tcactttcag	gcaatgttgt	ggggctgttt	aaacaacccc	aaagaaaacc	3060
atttggccaa	actgttagtt	tccaaacatt	ttacttcctt	ggtgtttaaa	taaattccta	3120
ccaagactct	gtagctggtc	ccagggaagg	agttggcctc	tcttctttat	agcccgccac	3180
agtcagtccc	ctgcacctgc	ccctcccaac	cccaggcctg	cttccccgtg	gccatggctg	3240
ctgcccggac	ctctctacac	acagaacctc	ctggaggcca	gctgtgggca	ccagccttgg	3300
cagggctgtg	gctgagccca	ggctgctggt	actctctctg	cagctgctcc	ctgctggcct	3360
ggctggacag	cgtccccacc	accactgggg	tcacctctgt	gctggtcaca	gctcactcag	3420
accttcaggc	aaatgggttg	gatcctgcct	ctctcccagg	tgtctcagtc	tctgcaaaac	3480
tcaaaaacct	cagaggcctt	gcagcctgag	gggtgtcaga	gacacctcct	tcgaatcagt	3540
aaacacctac	agattcacc	cagcagtgaa	aggactgctt	cgccacagag	gtttgattta	3600
ctcctaagta	attggaagg	atgccgagaa	taggttcctc	atgggtgggac	tagaggccct	3660
ctgctgacct	agttaacaga	gggctagggc	tgggtgtgct	cagccctga	aggttctagg	3720

cccatttggg	acaccccgcc	agaacctgcc	acaacctgcc	atgtggtgac	agctacctaa	3780
atcccagagg	ctcttgagct	ggagagcaga	cctctcaatc	tcagcaggcc	ccccacacag	3840
accccataac	cctagtctgc	cttcacagta	cagttcgtgg	ctatgtgttc	acggatgggtg	3900
ttgtttcacct	aaggtctctg	ccctgtgacc	ccaagggcgt	cctgagggca	gattccaagt	3960
ctgtttcgtc	cacccctcct	tccctagcag	cgggtccagg	gcctggcctg	aactagcttc	4020
ccacagagat	actgggtggga	tgatgaaggc	agccaggcgg	caagtgaaaa	acgcacttcc	4080
tgcatgtgct	ggctcctggg	attgaagtgt	ttgaggaagc	aaagtgaagt	gagctttcct	4140
cttgccggtg	tgtgtccttg	ggccgggagc	ctaccctctc	tgagcgttgg	ggtccttgct	4200
agtagaatgg	ggcatcctca	tagctcaagg	ggtggtgtgt	gaaaattgtg	ctattgtgtt	4260
actttaatga	tttttttttt	ttcgagacaa	agtctcacc	caacgcgcag	gctggagtgc	4320
agtggcgcca	tctcagctca	ttgcaacctc	tgctcctgg	gttcaagtga	ttctcctgcc	4380
tcagcctccc	aagtagctgg	aattacagga	gtgcgccacc	aggcccggca	tatttttcta	4440
tttttagtag	agagggggtt	ttaccatgtt	ggctaggctg	gtcttgaact	cctgacctca	4500
ggtgatccac	ctgcctcggc	ctcccaaagt	gctgggatta	caagcatgag	ccaccgcgcc	4560
cggcctactt	tagtgatttc	ttaggaggac	agagggaaacg	ggctggcaag	acaggcttgg	4620
aatgtgtttt	gggatcaagt	gccggtttct	gtctggcact	ggcgttctct	gtggggccat	4680
gatggacaca	ctgctgaggt	caagcgtgat	tcgtcttgcg	ctgtgcctgg	cagtctcatt	4740
ggaaagtctt	gtagacatcg	tgtggatggg	gctcttccc	gccaagccct	tggggacctt	4800
ccaggactgt	gatctcccca	cagtggctgt	taagcaggga	cctttcgtga	agtggagtct	4860
ctggtccctt	ccaagtcata	gctagacag	gactcgggca	tcgccaaagcc	tggctgatta	4920
ttcactggat	gaggagacag	gcccagagag	gggcaggaaac	ctgcccagg	tcacccagca	4980
ggccccagag	gtttcggctt	cggattctcc	ctgctcatcc	ctggatgtag	tgctgctgtg	5040
gatgtggttc	tgtgctgggg	gctgtggaga	gcagggggct	tgtgccagga	ccccagttag	5100
ggtggcgccc	tcgccatgag	gccgactgtt	ggtatggggc	ggccatccac	tggggtgtgg	5160
ggaggaacag	ctttcctgag	gaggaggtgg	cgggaggaac	agcttccctg	aggaggaggt	5220
ggcgggtgctg	tgtgacctgg	gccttgaagg	acagggtccat	tgtcaacaga	acattttggg	5280
agtggagcct	agagggagaa	aatttgttga	aattcagatt	ccccctcccc	taccaataca	5340
caccaaataca	gatgccctg	accagatcta	aatttggctc	tcagagattt	ccattgtagc	5400
tgggcacttg	gggaaccttc	taagtgtctg	ctctgcctct	ccccagcctg	cctgcctcag	5460
tttcccagc	cctgggccc	tgtcgctgtt	gccatcacgt	gggcgcctc	tagtggaggga	5520
atcagattat	gcactccggg	gcttggagca	ggagtcaggga	ggggctcctg	tctttccttg	5580
aaacgttgga	tgccgggatc	ctggaacagt	ctctgcattc	ctcctggcga	gaaccagagc	5640
ctgggcacag	gggacctctt	gttgtttgaa	ggctgcagcc	tggcagggca	ctcaggagat	5700
ctggcagttg	gctgcagggc	caggtctagg	ggccagggca	tcagggaggg	tctgggctgg	5760
ttcagccccg	ggcccccttg	cagattgtga	cctgggcccc	tgtgcagggg	catggccaca	5820
ggatgctggg	aggggtctct	gaccctgacc	ttcttggtct	tgtgcacctt	tgagaccaga	5880
aaggtctgga	acaaatgagt	agacgatgcc	ctaacctggg	gagggagcca	catcctgatc	5940
ccagcaacct	cgggaaggat	ctgtcaggat	tatggggcac	cctggggggc	ccaagtctgc	6000
atgggtctcc	acttgcaatt	tctgtaggaa	gctctgataa	atccaaactg	ggggctcctag	6060
gacacagtca	gaaatgctga	taccgttgtg	tgtggagcct	cgggccttgg	gggtcaggag	6120
catgtggagg	gtgggcccag	ggggttcaga	agagaatcct	gtaacccccc	accccccaaa	6180
ctgaagccca	cctgagggcc	atggctgaaa	ggttgggggg	tctccgtgcg	tctgtggag	6240
tgggtggtga	ggagtccctg	ggtttgcacg	cctctggggc	tgagcggcgg	gaccccgctc	6300
acagcggatc	cctgggccc	gttgctcaga	tgtctcaga	gtgttgctgt	ggccacggag	6360
ggagcctgag	ttaagcttct	cctgtgccgg	ttgtacgtg	tcaggtcaca	ctgggtgagtt	6420
aggcagggca	cagatgccc	gagcagagg	aactttcctt	ggggattcaa	cacgtgcaag	6480
tcttaggggc	tggcaaatacc	tgccctcagc	tagagagggg	gcttttatit	gagaccagaa	6540
tcacctgagc	atcctcctgt	ccccagctgt	gtccagcctg	tctgcaggga	catcctgaga	6600
ggaccaggct	ctccctcat	ccacctgcct	aagtgccact	ctgaaccctg	tccacctgtg	6660
ccgtggagg	gcgtgacctc	aagctgctca	gccagcagca	ggcttgggcc	tggggggcag	6720
cagagaccca	ggtggctgtg	gggtgggtgc	ttcgtggcgt	ggttctgaaa	cttcgttgga	6780
agtgtgtgga	cagtgccttg	cctgttctct	gtgggacctt	atttagaaac	gaggtctgag	6840
ttactggggg	tcactactgt	gttctgatgg	cccagctgtg	tggaggccgc	ggtgcagccc	6900
catccaagga	gccaggggccc	tgggtctagc	cgtgaccaga	atgcatgccc	cggaggtgtt	6960

tctcatctcg	cacctgtgtt	gcctgggtgtg	tcaagtggtc	gtgaaactct	gtgttagctc	7020
ttggtgttcc	tgaaagtgcc	cccgggtctc	aggcctcaga	accagggttt	cccttcatct	7080
cgggtggcctg	ggagcatctg	ggcagttgag	caaagagggc	gattcacttg	aaggatgtgt	7140
ctggccctgc	ctaggagccc	cccggcacgg	tgtctggggc	tgaagctgcc	ctcgggtggt	7200
ggagaggagg	gagcgatgaa	gtggcgctga	gctgggcagg	aagggtgagc	ccctgcaagg	7260
tgggcatgct	ggggacgctg	agcagcatgg	ccagcagctg	ggtctgcagc	ctggtacccg	7320
gcgggacttg	tggttggggc	tggtttgtgg	ccaggagagg	ggctggcagg	agacaagggg	7380
gactgtgagg	cagctcccac	ccagcagctg	aagcccaatg	gcctggctgt	gtggctctca	7440
gctgcgtgca	taacctctca	gtgcttcagt	tctctcattt	gtaaaatgag	gaaacaaaca	7500
gtgccagcct	cccagagggtg	tcatgaggat	gaacgagtga	ccatgtagca	tgggctgggt	7560
gcgtgtcacc	taacatcacc	agcctttgca	aggagagccc	tgggggcctg	gctgagtatt	7620
tcccttgccc	ggcccacccc	aggcctagac	ttgtgcctgc	tgcaggccct	tgacccctga	7680
ccccattgca	cctgtctcca	caggagccga	ggaggtgctg	ctgctggccc	ggcggacgga	7740
cctacggagg	atctcgctgg	acacgccgga	cttcaccgac	atcgtgctgc	aggtggacga	7800
catccggcac	gccattgcca	tcgactacga	cccgttagag	ggctatgtct	actggacaga	7860
tgacgaggtg	cgggccatcc	gcagggcgta	cctggacggg	tctggggcgc	agacgctggt	7920
caacaccgag	atcaacgacc	ccgatggcat	cgcggtcgac	tgggtggccc	gaaacctcta	7980
ctggaccgac	acgggcacgg	accgcacgca	ggtgacgcgc	ctcaacggca	cctcccgcaa	8040
gatcctggtg	tcggaggacc	tggacgagcc	ccgagccatc	gcactgcacc	ccgtgatggg	8100
gtaagacggg	cgggggctgg	ggcctggagc	cagggccagg	ccaagcacag	gcgagagggg	8160
gattgacctg	gacctgtcat	tctgggacac	tgtcttgcac	cagaaccggg	aggagggtct	8220
gttaaaacac	cggcagctgg	gccccacccc	cagagcgggtg	attcaggagc	tccagggcgg	8280
ggctgaagac	ttgggtttct	aacaagcacc	ccagtggctc	ggtgctgctg	ctgggtccat	8340
gcgtagaaag	ccctgnaaac	tggagggagc	cctttgtccc	cctgncttca	gtttccctcat	8400
ctgtagaatg	gaacggtcca	tctgggtgat	ttccaggatg	acagtagtga	cagtaagggc	8460
agcctctgtg	acactgacca	cagtacaggc	caggcctctt	tttttctttt	tttttttgag	8520
atggagtctc	actctgtcgc	ccaggctgga	gtgcagtggg	gtgatctcag	ctcactacaa	8580
cctctgcctc	ctgggctcaa	gtgattctcc	tgcctcagcc	tcctgagtag	ctgggattac	8640
aggtgcctgc	cactgtgctt	ggctaattgt	tgtatttttg	gtagagatgg	ggtttcaccg	8700
tcttggccag	gctggctcga	aactcctgac	ctcaggatgat	ccacctgcct	cagcctccca	8760
aagtgtctggg	attacaggca	tgagccacca	cgcccgggtca	ggccaggcct	cttttgaaca	8820
ctttgcacac	catgggtctt	ttcatccagg	gggtaggta	cagttgtaca	gttgaggaca	8880
ctgaagccca	gagaggctca	gggacttgcc	cagggtcaca	cagcaggatg	tggcagggtg	8940
ggggctgggg	ctggcagcgt	ggctccagct	ttccagcata	gaaatctgtg	aaagcagata	9000
gtttgtcggt	cggtagggga	gactttctga	gaccgcctcc	agcggctcag	agggtagtag	9060
ccaggggcct	tcctgggggc	tcataaccca	gaacactgaa	tgggaaaacc	ctgatggagg	9120
aggcgagtg	gagctgtggg	tgccgatggg	aagtcccaga	ggagctggga	ggtcagtagc	9180
ggtgctgccc	tctgtggagc	acttagtggg	caccagggtg	gtttccagg	tcattggccct	9240
gggacctgaa	gctcagaagg	tgaagtaact	tgcccagggc	accgctcggg	cagcggcggg	9300
cagaggattt	gtgggctgtg	gagcctgtgc	tcgtggccca	gccctggggg	ttgtgagtgt	9360
gctggccggg	gagcttttcc	tgcaagtgga	ctgggtgtcta	ggagccagca	tgtcaggcag	9420
caggcagcgg	gagtgcagca	ggcagcggga	gcacagcagg	cagagggcgg	ggctcgagca	9480
gccatccgtg	gacctgggg	cacggaggca	tgtgggagag	ggctgctcca	tggcagtggc	9540
tgaagggctg	ggttgtgccc	cgaggagggt	ggatgagggt	aagaagtggg	gtccccaggg	9600
gcttttagcaa	gaggaggccc	aggaactggt	tgccagctac	agtgaaggga	acacggccct	9660
gaggtcagga	gcttgggtcaa	gtcaactgtct	acatgggcct	cgggtgtcctc	atctgtgaaa	9720
aaggaaggga	tggggaagct	gactccaagg	cccctcctag	ccctggtttc	atgagtctga	9780
ggatcccagg	gacatgggct	tggcagctctg	acctgtgagg	tcgtgggggtc	caggaggggg	9840
caccgagctg	gaagcgggag	gcagaggggc	tggccggctg	ggtcagacac	agctgaagca	9900
gaggctgtga	cttggggcct	cagaaccttc	acccttgagc	tgccacccca	ggatctgggt	9960
tccctccttg	gggggccccca	gggaacaagt	cacctgtcct	ttgcataggg	gagcccttca	10020
gctatgtgca	gaaggttctg	ctctgcccct	tcctccctct	aggtgctcag	ctcctccagc	10080
ccactagtca	gatgtgaggc	tgccccagac	cctgggcagg	gtcatttctg	tccactgacc	10140
tttgggatgg	gagatgagct	cttggccccc	gagagtccaa	gggctgggtg	ggtgaaaccc	10200

gcacaggggtg	gaagtgggca	tccctgtccc	aggggagccc	ccagggactc	tggtcactgg	10260
gcttgccgct	ggcatgctca	gtcctccagc	acttactgac	accagcatct	actgacacca	10320
acattttacaa	acaccgacat	tgaccgacac	cgacatttac	cgacactgac	atttaccac	10380
actgttttacc	aacactgaca	tctactgaca	ctggcatcta	ccaacactga	cattttaccga	10440
cactgacatt	taccaacact	atttaccac	actgacatct	actgacattg	gcatctacca	10500
acaccaacat	ttaccgacac	caacatttac	caacactgaa	atttaccgac	accgacattt	10560
accgacaccg	tttaccacaa	ccgacgttta	ccgacaccga	cattttaccga	cactgatatt	10620
taccaacact	gacatctact	gacgctggca	tctactgaca	ccgatgccag	catctacca	10680
caccgacatt	taccaacact	gacattttacc	aacactgaca	tttaccgaca	ttgacattta	10740
ctgacactga	catctactga	cactggcatc	tactgacact	gacgtttacc	gacactagca	10800
tctactgaca	ctgacattta	ccaacaccag	catctacca	caccgacatt	taccaacact	10860
gacattttact	gacactgata	tctactgaca	ctggcatcta	ctgacaccaa	cattttacca	10920
caccagcatc	taccaacacc	gacattttacc	aacaccagca	tttaccacaa	ccgatgttta	10980
ccaacgccga	cgttttaccga	cgccagcatc	taccaacact	gacattttacc	gacaccgaca	11040
tttaccgaca	ctgacattta	ctgacactga	catctactga	tactggcatc	taccgacact	11100
gatattttacc	aacgccagca	tctactgaca	ctgatgttta	ccaacaccga	cattttacgag	11160
caccgacatt	tactgacacc	aatattttact	gacatcaaca	tttagccatg	tgatggggggc	11220
cggtctgggg	gcaggccttg	ctcttggcac	tggggatgct	gcagagacca	gacagactca	11280
tgggggtcatg	gacttctgct	tcttctccag	cctcatgtac	tggacagact	ggggagagaa	11340
ccctaaaatc	gagtgtgcca	acttggatgg	gcaggagcgg	cgtgtgctgg	tcaatgcctc	11400
cctcggtgg	cccaacggcc	tggccctgga	cctgcaggag	gggaagctct	actggggaga	11460
cgccaagaca	gacaagatcg	aggtgaggct	cctgtggaca	tgtttgatcc	aggaggccag	11520
gcccagccac	cccctgcagc	cagatgtacg	tattggcgag	gcaccgatgg	gtgcctgtgc	11580
tctgctatct	ggccacatgg	aatgcttgag	aaaatagtta	caatactttc	tgacaaaaac	11640
gccttgagag	ggtagcgcta	tacaacgtcc	tgtggttacg	taagatgtta	tcattcggcc	11700
aggtgcctgt	agacacagct	acttggagac	tgaggtggga	ggatcgctgg	agtccaagag	11760
tttgaggcca	gcccgggcaa	aggggacaca	ggaatcctct	gcactgcttt	tgccacttac	11820
tgtgagatct	aaattatttc	acaatacaaa	attaagacaa	aaagttaatc	acatatccac	11880
tgccctgctt	aagacagaaa	acatgggtgt	tgttgaagcc	agaggcagct	gctggcctga	11940
gtttggtgat	tggttcctaa	gcagttgaag	gcagttttgt	ttttccatag	atgtctgttc	12000
tccctttgct	gggtgcagcc	tcgccctgct	gctgtggtcg	ggtttcagtg	gcctcgctcc	12060
gtggacgcag	cctcgccctg	ccgctgtggt	cgggtttcag	tggcctcgtc	ccgtggacgc	12120
agcctcgccc	tgccgctgtg	gtcgggtttc	agtggcctcg	tcccgtggac	gcagcctcgc	12180
cctgcccgtg	tggtcgggtt	tcagtggcct	cgccccgtgg	acgcagcctc	gccctgccgc	12240
tggtgctggg	tttcagtggc	ctcgtcctgt	ggacgcagcc	tcgccctgcc	gctgtggtcg	12300
ggtttcagtg	gcctcgctcc	atgggcgtgc	tttggcagct	ttttgctcac	ctgtggagcc	12360
tctcttgagc	ttttttgttt	gttgtttgtt	tttgtttgat	tttgtttgat	tgtttggttt	12420
tggtgtcggt	gttgttgccc	aggctggagt	gcagtggcgc	gatctcagct	cactgaaacc	12480
tctgctcctc	tgggttcag	ccattctcct	gcctcagcct	cccacatagc	tgggattaca	12540
agtgcgccgc	accacgcctg	gctaaaattt	gtatttttag	tagacagggg	gtttcaccat	12600
gttggtcagg	ctggtctgga	actcctggtc	tcacatgatc	cacctgcctc	ggcctcccaa	12660
agtgttggga	ttacaggcgt	gagccaccgc	gcccagccct	ctgttgagca	tattttgagg	12720
ttctcttggg	gccagtgata	tgtacatgtg	tccccatcgc	accatcgta	cccattgagg	12780
tgacattggg	gcctctcctc	gggttggatg	cctccctctg	tttccagcaa	cttctgaagg	12840
attttcctga	gctgcatcag	tccttgttga	cgtaaccatc	gggttcacct	ttgctctcct	12900
cagggctccc	aggggaggcc	cgaatcaggc	agcttgcagg	gcagggcagg	atggagaaca	12960
cgagtgtgtg	tctgtgttgc	aggatttcag	accctgcttc	tgagcgggag	gagtttcagc	13020
accttcaggg	tggggaaccc	agggatgggg	gaggctgagt	ggacgcctct	cccacgaaaa	13080
ccctaggagc	tgcaagggtg	gccatttcct	gctggagctc	cttgtaaatg	ttttgttttt	13140
ggcaaggccc	atgtttgcgg	gccgctgagg	atgatttgcc	ttcacgcata	cccgtacccc	13200
gtgggagcag	gtcaggggact	cgctgtctg	tggcacacca	ggcctgtgac	aggcgttgtt	13260
ccatgtactg	tctcagcagt	ggttttcttg	agacaggggtc	tcgctcgctc	accaggcgca	13320
gagtgcagtg	gcgcaatcac	ggctcgctgt	agcctcaatc	tccctgggct	caggtgatcc	13380
tcctgcctca	ccctctgagt	agctgggact	acagacacat	accaccacac	ccagctagtt	13440

tttgtgtatt	ttttgtgggg	ggagatgggg	tttcgctgtg	gtgcccgaagc	tgatctcaaa	13500
ctcctgaggc	acaagcgatc	cacctgcctc	ggcctoccaa	agtgtctggga	tgacaggcat	13560
cagccgtcac	acgcagctca	atgattttat	tgttggtaaaa	taaacatagc	acaaaattga	13620
tgatttttaac	catttttaaag	tgaacagttc	aggctggggc	tggtggctta	tgcttghtaat	13680
cccagctactt	tgagaggctg	aggtgggcag	atcacctgag	gtcaggagtt	tgagaccagc	13740
ctggccaaca	tgatgaaatc	cagtctctac	taaaaataca	aaaattagcc	gggcatgggtg	13800
gcagggtgcct	gtaatcccag	ctactcggga	ggctgaggca	ggagaatcgc	ttgagcccgg	13860
gaggtggagg	ttgcagtgat	ctgagatcat	gccactgcac	tccaatctgt	gtgacagagc	13920
aagactctgt	cttgaaaaat	aaataaataa	aaaaaatttt	aaaaagtga	caattcaggg	13980
catttagtat	gaggacaatg	tggtgcagg	atctctgcta	ctatctactt	ctagaacact	14040
ttcttctgcc	ctgaaggaaa	ccccatgcc	accggcactc	acgcccattc	tcccctctct	14100
cccagcctct	gtcaaccact	aatctacttt	ctgtctctgg	gggttcactt	cttctggacg	14160
ttttgtgtga	ctggaatcct	gcaatatgtg	gtccctgcgt	gtggcttctt	tccatagcat	14220
tgtgttttcc	agattcaccc	acacattgtc	gcacgttatc	agaatctcat	tcctgactgg	14280
gtgcagtggtg	ttaggcctgt	aatcctaaca	ttctggggagg	ccaaggcggg	acgatcactt	14340
gaggcaggag	tttgagacca	gcctggccag	cctagcaaga	ccccagctac	caaaaaattt	14400
taaaagttaa	ctgaacgtgg	tggtgggtggg	cacttgtggt	tcccagctac	ctgggaggct	14460
gaggttggag	gatcgcttaa	gcccaggagg	tcaaggctgc	agtgagctat	gatcgacca	14520
ctgcactcca	gcctggacaa	cagagcaaga	ccctgtctga	aaaaaaaaac	aaaaaaaaaa	14580
gttcctttct	ttttgtggct	ggatgacatc	ccattgtatg	gccacagcac	attttgtttg	14640
tctgtttatc	gggtgggtggg	cagtgggttc	caccttttgt	ctcctgtgaa	taatgtctgt	14700
gtgaacattt	gaattcaagt	ttttgtttga	acacctgttg	tgaattattt	ggatatatgt	14760
gtaggggtag	gattgctgag	tcctatggta	atgttaagg	tgacttactg	aggaaccatt	14820
aaactgtttt	caacagtggc	tgcgccgttc	tgcacccca	ccggcagtg	gtgagggttc	14880
tgactttacc	tcctcacaaa	cgttctttt	ccatttaaaa	aaatattcag	ccagggtgctc	14940
tggtctcacgc	ctgtaatccc	agcacttttg	gaggccgtgg	cgggcggatc	acctgagggtc	15000
aggagttcga	gacgagcctg	gccaacatgg	tgttaaccca	tctctaccaa	aaatataaaa	15060
attagccggg	tgtggcagcg	ggcgccgtga	atcccagcta	cttgggaggc	tgaggcagga	15120
gaatcacttg	aacccgggag	gcagaggttg	cagtgaacca	agatcgcgcc	actacactcc	15180
agcctgggtg	acaagagtga	aactccatct	aaaataaaa	aaaaataaaa	ataataaaa	15240
atttattaaa	acattcatca	cagccagcct	agtgggtgtc	ccatgtggct	ttgcctcgca	15300
tttccctgat	aactaggatg	ctgagcgtct	tgtcccaggc	ttgccacacc	tcagcacttt	15360
gagatacgtc	gcacagtccc	catttgcgaa	cgagaaatga	ggtttaggga	acagcagctg	15420
tgtcatgtca	cacagcgagc	aggggggtctc	tgagccgtct	gacccacag	ccgaccaagc	15480
tccaatcctt	accgcctcct	agtgttgtgg	atgtagccca	gggtgctccc	acatttttca	15540
gatgagaaca	ccgaagctca	aaacaggagc	gttttgtcca	cattggatac	acgatgtctg	15600
tggtttgggtc	ctgaagtcac	tttatatctc	agtgggtccag	actggagtag	gacaggggggt	15660
tctggggaat	ggggaagggtg	tctcagggtga	aaggaaggaa	ttccagattc	tccatactgt	15720
ccttgggaag	ttagaagact	cagagggtct	ggcaaagtca	gacaaagcaa	gagaaatgca	15780
gtcaggagga	agcggagctg	tccaggaaca	gggggggtcgc	aggagctcac	ccccaggaa	15840
tacacttgct	ggggccttgc	tgtcacaatg	acgtgagcac	tgcgtgttga	ttaccactt	15900
tttttttttt	tttgagggtg	agtctcgctc	tcttgcccag	tctggagtgc	agtggcacga	15960
tctcggtcca	ctgcaagctc	tgcctcccgg	gttcatgcca	ttctcctgcc	tcagcctccc	16020
gcgtagctgg	gactacaggc	gcctgccacc	gcgcccggct	aatttttgtg	tttttagtag	16080
agatgggatt	tcactacatt	agccaggatg	gtctcgatct	cctgacctca	tgatccgccc	16140
gtctcggtct	cccaaagtgc	tgggattaca	ggcgtgagcc	accgcgccc	gcccgatctc	16200
ccactttaag	aatctgtctg	tacatcctca	aagccctata	cacagtgtctg	ggttgctata	16260
gggaatatga	ggcttacagg	ccatgggtgt	ggacacacag	aaggacagga	ggtcaggagg	16320
tagaagggtg	gagagaggga	acaggcggag	gtcacatcct	tggctttcaa	aatgggcccag	16380
ggagagacac	cctctgagca	tggtaggaca	ggaaagcaag	attggaacac	attgagagca	16440
accgaggtgg	ctgggcgtgg	tggcttacgc	ctgtaatccc	aacacttttg	aaagctgagg	16500
tgggtggatt	gcttgaggcc	aggagttcaa	gaccagcctg	gccaacatgg	tgagacccc	16560
tctctactaa	atatacaaaa	attagccagg	cgtgatgggtg	catacctgta	atcccagctg	16620
cttgggaggc	tgaggcagga	gaattgctta	aacctgggag	gcggagggtg	cagtgagccg	16680

agatcccccc	actgcactcc	agcctggggc	acagagtggg	actccatctc	aaaaaaaaaa	16740
aaaaaaaaaga	taaaaagacc	aaccgaggaa	ttgaagtggg	ggggcgtcac	agtagcagaa	16800
gggggatcgt	ggagcaggcc	accctgtggt	catgcactgg	aagctcatta	cctgacgatt	16860
tggagctcat	cactgggggg	ctaaggagaa	tagatactga	aggatgagga	gtgatggcgc	16920
ggggcacggg	tgtcttttgt	ggccagaact	tggggactgc	tggggtgcct	cactgcaggc	16980
cttctcagcg	ccctttatat	gcttacacag	gctgtttcta	agagggggat	acattgcata	17040
agcgttttca	gactacctca	tcatgggtcc	ctttctttac	cctctgtggc	cctgggtggc	17100
cactctctgg	gaaggtgcag	gtggatgccc	agacccgccc	tgccatccac	ctgcacgtcc	17160
agagctgact	tagcctcgag	attgctgctg	gcacctcctg	ccccgggaca	cctcggatgt	17220
gcccgtggag	atgctggctc	tgtgttttct	gctggagtgt	ggtgcgtctt	ttcctcctgc	17280
aagtggccac	cgctcttggg	tatgtcctca	ggcttctgcg	agtcatggct	gcttctcagg	17340
tccttgccca	gcgccaggag	caaaccctcc	tggcactttg	ttcaggggtg	gatgcgccag	17400
tgttcctgct	gtggaccgcc	atctcacatg	agggctcttg	gcctgcaggc	tcgttcagga	17460
aacacccgct	gagtatgcag	tgtgtgccag	ctgtgtccca	ggcaatggcg	gggacagtgg	17520
ctgctgctgg	ggttgtgggt	gcttctgggg	actctgggga	cagctgaggt	gcaaggagcc	17580
acggctcctt	gaggatgcag	ttggactcca	ggtggaagg	atggttgggg	gaggtataaa	17640
tggggtcagg	gaggagacac	atttgggaaca	atgggaacat	ttttaagatg	ctatgtcggg	17700
aggcaacaag	gtggccaacc	caggtgctga	ggagcccaca	ccagccctgg	acgtgttttg	17760
ccgctcacct	ttgctgggga	gtggtgggag	agaggattcc	gttccacgtg	gtggtgtgcg	17820
cagctgggct	gtgtggagct	gggcgctagg	aggaagggtc	tttctgcggg	gctagccggg	17880
ctctgccttt	gaacacaatc	aggctccagg	ttttcagcat	ccagtgcatt	agaggacttc	17940
acgggcagct	gtggctgatc	ccttgatgaa	ttgggagaag	aacaaagggtc	tatgaaatga	18000
ggtttcatgt	agatggcatt	agagacgccc	acaacagatt	tacagagtgg	agcggagacg	18060
gcggatgggt	ctgggaggcc	cctcctgctg	gccttgactg	tgacagctgt	cctgggaatc	18120
agcttccagg	ccgccccagc	agcctgactg	acacacacag	gggttttagc	cccatcctgc	18180
gaccagctgt	tgccatcatc	agtgcagct	gggagtggcg	gtggttccag	ccctgggcac	18240
cctccccacc	tgctggggcc	caccagggc	agtcccgaca	cctacagggt	gcttggagcc	18300
gcatccgagt	cctgccccac	cacgtgtgaa	gcccagtggt	tcgtgggctg	aggtcccctg	18360
attgcatccc	cacttccctt	ctgcttcaca	tagctgcctc	ttctcaccgt	ttttccagcc	18420
tcctgggcta	ggaattccag	tgttgtgctg	gctttgcccc	aggacacctc	cttagccctc	18480
ttctgagtc	tagagccccg	ggggttgga	gtcctggccc	ctgggacacc	tgacgccaca	18540
ctcagcttct	cctgtgagcc	tccagcatgt	ccctcaggga	ccaagccctc	acgttcttgc	18600
ctccccgccc	acctgggctc	agccagggga	aggcctgggt	gggagcgtct	cccctctgcc	18660
ctgcccttct	cccctcctac	cctgcccttc	tctcctctgc	cccgccatgg	cttttatatc	18720
ctgtgccaca	agacatggct	gtgtgtgaaa	gtggcagggt	ctggcatctc	tgtgggtctc	18780
tgaggccccc	gctccagtgc	cactcttccc	acccgctggc	cgtgccctca	tgctggaggg	18840
acagcccagc	cctctcccga	accccagccc	catgtgcccc	gctgcccccg	gccctctccc	18900
ctggaagccg	gggtcactcc	agccgtatgc	catggtgggg	acatcctgct	tccttggcct	18960
tccagggaag	gtcctctttc	caaatggcga	cacctgggtc	ctgcctggag	gctggaagct	19020
gtggcccttg	tatgcccctc	cagggtctgt	gcgctcgggt	ggcccaggtt	cccatcaccg	19080
tcacatcatc	catcatcatt	gtcatttcgc	ttgtctgtga	gccggcctgg	tctcccagag	19140
cagagaccct	ctgaggtcca	gcctgagttg	gggtctccgt	gctgaccctc	gacggggact	19200
caggacgtac	caggtctggg	tcaggagtga	cccccaaacc	tcgtgccctt	tgacaggcac	19260
ccctgacttt	tgctaagtgg	gtggagggtga	catcacttac	agcgggagtg	atgggacagg	19320
gtctgttggc	tgactgtgct	tcccagggat	ctggggagag	gctatatccc	tgggcttttg	19380
cactgcagag	ctgtgtgtgt	ttgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	19440
gtgtgtgttt	gcgtgcgcgc	acatgtgtat	aagatctttt	tttattacat	gaagcaagat	19500
aactgttgct	gtttcctttt	gggttttgtg	ttcaacagag	tgggttactt	cttccctcag	19560
acaacagaa	tctccccctt	aaacacgtgc	tgctcagagg	tgggtcttgg	gctcatgtct	19620
gtttgcacag	ccgagtcaga	ggaaacacag	ggttcttcat	aaaaaactg	cacagcaggc	19680
gactgtccag	agtcagcctg	caggacggca	gcagccctgc	ccctcagagc	acagctaggg	19740
tgggctgctt	tgggatctcc	cgctattccc	tcccagctgg	cagccggcgg	ccggcccatt	19800
ccttggtgtg	ctgggtcagg	gggcgtgcgc	ctgctctgct	caccctggga	atgggacaga	19860
agctggcagc	tcggagagga	cagggtctga	cccttgggtg	gcctctgggt	ggaccatctc	19920

attgtcctca	gacacagcct	ctcgggtcta	gtttcatttc	ctgaaaaaca	agtgcacaga	19980
actagagcag	gagtcgagag	ctacggcccc	cgggccagat	ccagccctgc	cacctgtttt	20040
cacaccatgc	tcaagctgag	tgggtttttac	atTTTTtaat	tacttgaaaa	aaaaaaagcc	20100
aaaggagggtt	tcatgaccca	tgaaaattat	atgggaattca	aaaaaaaaaa	attatatgga	20160
attcaaattt	cagtgtccat	aaataatttc	ttgagacagg	gtctcgtct	gtcaccagc	20220
ctggagtgc	gtgctatggc	atggctcgct	gtacccttga	cctcccaggc	tcaagcgatc	20280
ctcctgtctc	agcctcctga	gtagctggga	ctaoggggtg	gtgccacca	gcccggctaa	20340
ttttttttta	atttttagtaa	agacagggtc	tttctatgtt	gcccaggtt	ttctggaact	20400
ccatcttggc	ctcccaaagt	gctgggatta	caggctcgag	ccacggagcc	cagcctgttt	20460
ttgttttttc	actgataaag	ttttgccggg	tgtggtagtg	tgtgcctcta	gcgatttggg	20520
aggctgaggt	gggaggatcg	cttaagccca	ggagtttgag	gctgggctca	agtgatcagg	20580
agggtgaacta	tgatcatgtc	attgcattcc	agcctgggtg	acagagcaag	aacctatctc	20640
ttaaaaatat	atatttaaaa	agtattgggt	gtgggtggctc	acgcctgtgg	ttccagctac	20700
ttaggcatct	gagggtgggag	gatggcttga	gcccaggagt	ttgaggttgc	agcgagccaa	20760
gatcgtgtca	ctacactcta	gcctgggtga	cagagccag	accctgcctc	tttaaaaaaa	20820
aaaaccaaaa	aacatgtatt	ggaacacagc	catgcctgtt	cagtcacgtg	ctctccatgc	20880
tgctttctgc	tccagagacc	cttatggcct	gaaagctgaa	aataTTTTct	atcctttaca	20940
aaaaagtttg	ctgacctctg	tcctggaaaa	ttcatctccc	aagttctctt	ccggcactgg	21000
cgttcctggg	tgtcctaaat	ttggcccttg	ttattttctga	actctgtttt	ggctctgttc	21060
cctcccagga	gccaggacag	gcacgtttct	tgcattctgt	cccctgacgc	ccagaggctt	21120
ggctcggctc	aggcattctt	ggaaatatct	ggctccagga	aaggcagagg	cctcctgagt	21180
cggcccagag	ggaacctgcc	ccaggctctg	gggaggcctg	acccagcaga	gtggcctttg	21240
ccgatgggtt	gggcccgtca	agatgtgctg	aaagtgttcc	tcagaaggcc	actttgggat	21300
tccttcctcc	agatttagag	caactgagag	ctgctcattg	caagcctgat	gttttcccag	21360
ttggccgggt	ccaccgggtg	ccctgggatt	ctgggatctg	ggtggaaagt	agggggcttg	21420
ggggagtgtc	ctgggttctg	gaatccaggt	ggcaagtggg	gaggttcagg	gagtggcttc	21480
tgagccacca	taggggtctc	tgtgggaggc	tctgcccac	caggagattc	cgcaggccct	21540
gccggccag	agccagcgtc	ttgcgcttgc	cgaggctaca	gccagcccca	gccgggtgga	21600
acagcccgtc	gcctcctctc	actttgtttt	ggggccacct	gggagtgtgg	agcaagggtg	21660
gagagggagg	aagtggctgc	cggccgctgc	ccagcacctt	tgtttgcctt	gggccctctg	21720
tgggtcctt	tttattgtct	ttcaatgaag	ccagggaat	ggacttctt	gcctcacttc	21780
agttcaacat	gtctggaagt	ttggtattaa	aattaagaaa	gtgtggaaat	agagcaagaa	21840
gagaaaaatc	tctccaagag	ataatagtga	cctctgagct	gggcgcgggtg	gctcacgcct	21900
gtaaatccca	gtactttggg	aggctgaggc	gggcagatca	cctgagggtcg	ggagtttgtg	21960
accggcctga	ccaagatgga	gaaacccctg	ctctactaaa	aataaaataa	taaataaata	22020
aataaaatac	aaattagcca	ggcatgggtg	cgcttgccta	taatcccagc	taaggcagga	22080
gaatcgcttg	aacctgggag	gcaaaggttg	cagtgaagca	agatcacgcc	attgcactct	22140
agtctgggca	acaagagtga	aactccgtct	caaaaaaaat	aaataaataa	aaaataaaaa	22200
tagtgacctc	tggccagggtg	tggcagctca	taccgtaat	cccagcactt	tgggaaggag	22260
gccgagatgg	gcagattgct	ttagcacagg	agtttgagac	cagcctggcc	aacatgggtg	22320
aaccccatct	ctacaaaaat	agaataaaaat	ttaagaggta	atagtgcctt	tttggtagat	22380
cgaacactgg	attgctttct	ttttctaaat	gctgattctt	ttctttgtgg	tgtttgtgtt	22440
ctgtgccgat	gtccctcccc	cagccctgtt	attgtgagtg	gaagaagggg	aaagggttcg	22500
cccgtactg	tgagcccttc	ctctcacgct	gggtgtcctt	ggagaagcct	gcacttcttc	22560
attgtacgcc	agggctgggt	ccctccctgg	agtggttctg	tgtgtctggg	atggggccaa	22620
cccctcagat	gttttctgag	tgtcacacac	aggtgtgtgc	attcatggcc	tttgcgtgtc	22680
ttcctgttgt	ggaggcaaaa	atgtgaagaa	ccctagatga	ttttgggacc	agggtcccat	22740
cacctgctgt	tcattgcaca	cggagcatc	caggcatggg	tggagagctc	agacttccag	22800
gcacggctgc	aggggtctgt	ctaaccatgt	tcccgcctgc	ctgctcgtca	gaaccgcctg	22860
ttgggagctg	ttatcatgat	accatacctg	ggccctgggc	tatccgattc	tgacttaatt	22920
gctccaggtt	ggggccaggc	cgttgtttgc	tgttttgttg	tttcttctgt	gacgttagcc	22980
actgggctaa	tctgagcccc	tcagttacag	gtggagaaac	tgagacccat	gggggtgcaa	23040
ggacttgccg	aggaccagga	gccccttggg	ggcagagctg	aggcggggcc	tggctttggg	23100
tcccagagct	tccagtcctc	ttcccgtctc	cctaacagct	tttttttttg	agacaagatc	23160

tcaccctgtc	accaggctg	gagtgcaatg	gcatgatctc	ggctcactgc	aatcttcgct	23220
agctgcgttc	cagcgattct	cctgcctcag	cctcccgagc	agctgggatt	acagggtgtgt	23280
gccgccatgc	ccagctcgtt	tttttttcta	cttttagtag	agatagggtt	tcaccatgtt	23340
ggccaggctg	atctcgaact	cctgacctca	aatgatccgc	ctgcctcggc	ctcccaaagt	23400
gctaggatta	caggctggga	tcacactgtg	cctggcccta	gcagctttgt	cctgtgccat	23460
ccaacaacag	atgaccgaag	tctttgtttc	ttaacatgca	ttccatctgc	cttacagttt	23520
tgccacctgc	aaaacagagg	acttgtcgct	tttctggtaa	gctggaaatg	taatctggta	23580
gcaggaggcc	tgtggaagct	tgcctttaat	ggccttggtg	ctctttcatc	ctgtcctgag	23640
agccggagaa	cttggaatgt	gcacctaaact	caaccttcct	gttaacatac	agtctctgcag	23700
gctcatggat	catcagaacc	acgtcctatc	tcacgcggct	gtatgcttcc	gttggttcag	23760
gtgtttttac	cttgacagta	ttttctcctc	ggtggctttt	gcggtgggtg	cttttaaatca	23820
gcattgactc	ttcaagaaaa	atatttagct	gctacatctc	agaggagaca	gggtggaaaag	23880
catctgagac	ctgcaggctc	agacttagaa	ccagaagtgc	cctcagagtt	catccggccc	23940
tgaccagcg	ggaaatgagt	tcacagagaa	gcggggagaa	tttgccccag	gccctgccgt	24000
tgctcataac	tgccccaggt	ccttacatct	gctccaggct	ctgccccagg	ccctgcagtt	24060
gctcataact	gccccaggct	cttatatttg	ctccaggctc	tgccccagggt	cctgcagttg	24120
ctctgtgtgg	tgggtgtgat	ctggagccct	ccgcccattg	ctgcacctgg	ggcaggcatt	24180
gctaattgat	cccaggactc	cttcctgcgg	agcacgccct	ggttctccag	gcagccgctg	24240
cctgtcagcc	tgcagtgggt	cgggagagga	cacctgcttg	cctggtctgt	tccaaatctt	24300
gcttctcatc	ccagcacagg	taggggggtg	tatgggaaag	ggatcctcag	ttggccctgt	24360
cactgctcta	tcagctgggg	acgtggcatc	ctagtgaaaa	catcatggcc	gggcgcgggtg	24420
gctcacgcct	ggaatcccag	cactttggga	ggctgaggag	ggtggatcac	ttgaggctcag	24480
aagtctgaga	ccagcctggt	caacatgggt	aaacccatct	ctactaaaaa	tacaaaaatt	24540
cgccagggtg	ggtggcgggt	acctgtaatc	cgagctactc	gggaggctga	ggcaggagaa	24600
tcgcttgaa	ctgggagggt	gagcttgcat	tgagccgaga	tcttgccact	gcactccagc	24660
ctgggcaaca	gagtgagacg	ctgtctcaaa	atctcaaaaa	aacaaaacaa	caaaaaacaa	24720
acaaacaaag	cgtcatttat	ccagcacccc	tggggaacca	tgctacctgg	tgttttatgg	24780
tacctggcaa	ggtgcagggt	aagtgtctgc	tcttgggcat	tgaacccgtc	ttgtttgggg	24840
cagctcaggc	cccaggcagg	gtccgggttg	gctctcggtg	gtgtggccct	ggcccatcca	24900
gacctatatt	tctgccgtcc	tgccagggtg	caatgttgat	gggacgaaga	ggcggaccct	24960
cctggaggac	aagctcccgc	acattttcgg	gttcacgctg	ctgggggact	tcatctactg	25020
gactgactgg	cagcgccgca	gcacgcagcg	ggtgcacaa	gtcaaggcca	gccgggacgt	25080
catcattgac	cagctgccc	acctgatggg	gctcaaaagt	gtgaatgtgg	ccaaggctcg	25140
cggtagagtc	ggggggtccc	aagccatggc	tcagccatgc	agacttgcat	gaggaggaa	25200
tgacgggtcc	atgcctgggc	ataagtgttg	agctcagggt	ccccgacctg	gggaaggggca	25260
ggacaggaaa	ggtgacagta	tctggccaag	gacagatggg	aagggaacaa	gggagctgat	25320
tagggagtg	ttatggacta	ggaatgtcgg	taacaatggt	tagaaaagtga	ctaacatttg	25380
ttgagcacct	gctgtgtgcc	cggccctggc	cgggagcctt	cgtgcccaca	gtgaccccg	25440
ctgcaaatgt	agttccttgc	cctactcgca	ctggggagca	ggacgcagag	ccgtgcaact	25500
cacagggtgc	aagctcagga	ctccctcctg	ggtctgcctg	ggctgggctg	tgcttgttgc	25560
ccctgtggcc	cacgcatgtg	caccttccac	ctgaaaagca	ggatcttcag	gacgctcccc	25620
gaggaggtcg	ttgtctggca	caatgatttg	tctcttcctg	aaaagggtgac	agagttacac	25680
tgagagagag	agcatccagg	tgccgagagg	acaggccctg	ggctcgcggg	cagggactct	25740
gtgtcctgcc	ggggtcccac	actgcacctg	cttgtcagag	gcactcagtc	aatctttgct	25800
gatgaaggat	gagaggacag	aggacgtgat	gcttgctgct	gcattgcctg	cagtcctggg	25860
tgagatgccc	gggttgactc	tgctgcccgt	cgggtggatg	tgatgtcaga	tccccggctt	25920
taaaatacga	gggagctggg	aattgaggga	gcagggtggg	gcagaaagca	cagccccgtg	25980
gaagcctgga	gctgaggcag	tgtgggcgac	ccctggagca	gtgagtgcct	ccttcattggc	26040
cttcacgcga	ccctgcagtc	ctcatgtagg	ggatgccatc	catgaattta	gttttcccag	26100
cctcctttaa	aaacgcgttc	atgctggggc	cggggcagtg	cagtggctca	catctgaaat	26160
cccaccactt	tgggaggccg	aggcgggtgg	atcatgagg	caggagatcg	agaccatcct	26220
ggctaacaag	gtgaaacccc	gtctctacta	aaaatacaaaa	aaattagccg	ggtgcgggtg	26280
cgggcgccctg	tagtcccagc	tactcgggag	gctgaggcag	gagaatggcg	tgaacccggg	26340
aagcggagct	tgcatgagc	cgagattgct	ccactgcagt	ccgcagtcgg	gcctggggca	26400

cagagcgaga	ctccgtctca	aaaaaaaaaa	aaaaagtaca	aaaaaaaaaa	aattagctctg	26460
ggtgtggtat	cacgcgccta	taatctcact	actcgagagg	ctgaggcgga	gaattgcttg	26520
aaccaggag	gtagagggtt	tagtgagccc	gtatcgtagc	actgccctcc	acctgggcaa	26580
tagagcgaga	ctctgtctca	aaaagaaaaa	aaaaaaaaga	acatttatgc	caggtgtggt	26640
ggctcatgcc	tgaaatccca	gaacttttga	agactgaggg	aggaggatca	cttgagccca	26700
gaaatttgag	agtgtcttcc	ctgggcaaca	tagagagacc	tcatctctac	cagaaaaaaa	26760
aaaatttagcc	cggcatgggt	gcataatccct	gtggtcccag	ctacttaggg	ggctgacgtg	26820
gcaggatcac	ctgagtctgg	aggcagaggt	tgaagtgagc	tgagatcatg	ccactgcact	26880
ccagcctggg	tgacagacag	agaccctgtc	tcaaaaaaaa	aaaaaaaaaa	aagcattttac	26940
tatccaccat	ggaaggtgag	actgacctgt	gagtgattgt	tcaaagaaca	aaaaataaac	27000
cccagagata	agacaaaagg	gtgcctccat	gggggtgtga	tttaaagctg	agaaattggg	27060
cttcttcccc	ctcccccttc	accccggtgt	ttgctaaagg	agatgggaaa	aaggattctt	27120
tttttggtctg	aaatatttaa	cactaaatta	aagccaattt	taacagcact	ttggttgatg	27180
agtgaatta	acagactggc	caaaaaataa	cgaacggctc	gtactatgtg	aaaaagaggc	27240
agctttggcc	atgctggggc	aatgtgagtt	ttcaggggtg	ctgggaatgt	ctgtgaatcg	27300
gaggaagggc	ctagctggga	ctctcaggag	ccaaggccct	gaggggcaac	ttgcctggtc	27360
cctgcctga	ggcgttcact	gctttcttcc	tgggccagat	cacaggcccg	gaggctggac	27420
cactgggctg	gcactcttgc	cgagctgtct	cctgacttcc	tgaccatgct	cctttcagca	27480
gccttgctgc	acttttagttt	ccttgaatga	aaaatgggga	tgagaatagc	tcctacctcc	27540
aaggtgaatg	gagtgaattc	ggacaggtga	ctccctggga	ccagtgcctg	gcgcctgaca	27600
aggtccagtc	agagcccgca	ctgctgttac	tgataccctt	ggctgtacca	ggggagaact	27660
tggttgccat	tgccaggtgt	tctcccacca	ccccactac	tgtccctggt	tgatgtgtgg	27720
cgggaataaa	gctgtgcaca	ttggagcttt	tggcacatcc	tggctttcag	gtgaaagggtg	27780
cgtgtgtgtt	tgagggttta	gcctggccaa	cccagccatg	aggtcggacc	tgacctgggg	27840
gtgagtcttg	agctcggcac	ccctgagctg	tgtggctcac	ggcagcattc	attgtgtggc	27900
ttgggcccga	cccccttccc	tgtctgggctg	ttgatgttta	gactggagcc	tctgtgttcg	27960
cttccaggaa	ccaaccctgt	tgcggacagg	aacggggggg	gcagccacct	gtgcttctgc	28020
acaccccacg	caaccgggtg	tggctgcccc	atcggcctgg	agctgctgag	tgacatgaag	28080
acctgcatcg	tgcttgaggc	cttcttggtc	ttcaccagca	gagccgccat	ccacaggatc	28140
tccctcgaga	ccaataacaa	cgacgtggcc	atcccgtcca	cgggcgtcaa	ggaggcctca	28200
gccctggact	ttgatgtgtc	caacaaccac	atctactgga	cagacgtcag	cctgaaggta	28260
gcgtggggcca	gaacgtgcac	acaggcagcc	tttatgggaa	aaccttgccct	ctgttcctgc	28320
ctcaaaggct	tcagacactt	ttcttaaagc	actatcgtat	ttattgtaac	gcagttcaag	28380
ctaatcaaat	atgagcaagc	ctatttaaaa	aaaaaaaaga	tgattataat	gagcaagtcc	28440
ggtagacaca	cataagggct	tttgtgaaat	gcttgtgtga	atgtgaaata	tttgttgtcc	28500
ggtgagcttg	acttcagaca	cccccccac	tcccttgctg	gtgcccgttt	gctcagcaga	28560
ctctttcttc	atttatagtg	caaatgtaaa	catccaggac	aaatacagga	agactttttt	28620
tttttttttt	tgagacagag	tcttactctg	ttggccaggc	tggagtaccg	tagcgtgagc	28680
tcagctcact	gcaacctccg	cctcccaggt	tcaagcgatt	cttctgcctc	agcctcctga	28740
gtagctggga	ctacagacat	gcaccaccac	accagctaa	ttttttttat	atttttagta	28800
gagacaggtt	ttcatcatgt	tggccaggct	ggtcctgaac	tcctgacctc	aggggaacag	28860
acgggggttg	cctcccaaag	ggcggaataa	acaggggtga	gccaccgttc	ccggcctagg	28920
aaaacttttt	gccttctaaa	gaagagttta	gcaaactagt	ctgtgggctg	gccttctgat	28980
tctgtaaaaga	aagtttgatt	ggtggctggg	tgcggtggct	cacacctgta	atcccagcac	29040
tttgggaggc	cgaggtgggc	agatcacctg	aggtcgggag	ttcgagacca	gcctcaccaa	29100
cgtggagaaa	ccccgtctct	actaaaaata	caaaaaaaa	attaaccggg	catggcggcg	29160
cctgcctgta	atcgagcta	ctcaggaggc	tgaagcagga	gaattgcttg	aacctgggag	29220
gcggaggttg	tgggtgagctg	agatggcacc	attgcactcc	agcctgggca	acaaaagtga	29280
aactccgtct	cagaaaaaaa	aaagtgtgat	tgggtgaacc	aaagcgcat	tgtttatgga	29340
ttgtctgtgg	cagcttttgt	tctgccgaga	tgagttgtga	cagatctgta	tgggctctaa	29400
agcctaaaac	atgtgccatc	cgccccctta	cagaaaaagt	gtgctgacct	ctgttctaaa	29460
gtattggaca	actacaatgt	ttgctcattt	attattctat	gatttgtttt	ctgctttttg	29520
ttgttgttgt	tggtgttgag	atagggtttc	cctctgtcac	tcaggctgga	gtgcagtggg	29580
gtaatttcag	ctcactgcag	cctcgacctc	ctgggctcta	gtgatcctct	catctcagcc	29640

tccctagtag	ctgggactac	aggcacacac	caccactcct	ggctgatttt	tttttttttt	29700
tttttttttt	gtggagacag	ggtttccgca	tgttgcccag	gctggtttca	aactcctagg	29760
ctcaaacacc	cacctcagcc	tcccaaagt	ctgggattac	aggcgtgagc	caccatgccc	29820
agcctattct	actgtttgta	ttacatagct	ttaaaagatt	ttttatgact	ttaagtcaca	29880
agggttcttt	gtagaaaaaa	atatatatat	aggaaagtat	aaaaagaaag	taaaaattgt	29940
ccataacctc	tccagccaga	gacgaccgtt	gctgacacct	cagcatattg	cctttaagtc	30000
ttttttctct	aagatagcat	ttctcttcat	cacagtcata	tgctacgcag	aattctgtat	30060
cctgattttt	tcacttgaca	ttacaacagg	tatttgatgg	cgctgtgaca	aactccttgg	30120
cacaatcttt	taaatgtatg	aaatactcca	ctgcacagat	gtttgctttt	aggcttaact	30180
gttcttttat	tttgcgtgtg	ctggttacag	ccgggcacag	tggctcatgc	ctgtaatcac	30240
aacactttga	gaggggtgag	caggaggatc	acttgagccc	agaagtttga	gaccggcctg	30300
ggcaacatag	tgagacccca	tctctacaaa	aaactttttt	aataagtcgg	gcgtagtggt	30360
gcatagctgt	agtcccagcc	accaaggagg	ctgagttggg	aggattgctt	gagccccagg	30420
aggttgatgc	tgcatgacc	tgagattact	ccactgtact	ccaacctgag	cgacagagca	30480
agacttgtct	ggggaaaaaa	aaaaaaaaaa	tatatatata	tatatatata	tatatacata	30540
tatacataca	cgcacacaca	cataatataa	aaatatatat	ttataaatat	ataatatata	30600
atataaaaaat	atataatttat	aaataaaatt	tataaattat	atttataagt	aaatatataa	30660
tatataatat	aaaaatatat	attatataat	atataataaa	atataataa	taaaaatata	30720
tatttataaa	taatatataa	tacatactta	taagtataata	tttaaaatat	atgtaatgta	30780
tatttttttaa	tgtatgatat	ataatataca	ttataaaata	cacatttata	ttattttata	30840
taaaatatat	ataaaatctc	caagttgctt	tttccaaaaa	ggtgtcttgc	tgcatttcaa	30900
acattcatth	aaaaacttga	atgctggtga	tctggtccag	aatgtgttca	gtagctgctg	30960
ccagtggcca	agcatctcgg	gagatgtcta	caaaacacgc	tgggttctggc	ctggcgtggt	31020
ggctcacgcc	tgtaatctca	gcactttggg	aggctgaggc	agggtgatca	actgaggctc	31080
ggatttcgag	accagccttg	ccagcttggt	gaaaccccat	ctctactaat	aatacaaaaa	31140
aattagccag	gcgtggtggc	atgtgcctgt	aatcccacct	acttggggagg	ctaaggctgg	31200
agaatcgctt	gaacccaggg	ggcagaggtt	gcagtgagcc	gagatcgcac	cattgcactc	31260
caggctgggc	aagaagagcg	aaactccgtc	tcaaaaaaaa	aaaaaaagat	gctgggtcct	31320
aaaatgtggc	ccttttcctc	ctcacctgct	gccagaccat	cagccgcgcc	ttcatgaacg	31380
ggagctcggg	ggagcacgtg	gtggagtttg	gccttgacta	ccccgagggc	atggccgttg	31440
actggatggg	caagaacctc	tactgggccc	acactgggac	caacagaatc	gaagtggcgc	31500
ggctggacgg	gcagttccgg	caagtcctcg	tgtggaggga	cttggacaac	ccgaggtcgc	31560
tggccctgga	tcccaccaag	gggtaagtgt	ttgcctgtcc	cgtgcgtcct	tgtgttcacc	31620
togtatgaga	cagtgcgggg	gtgccaaact	ggcaagggtg	caggctgtcc	gtgtggccct	31680
cagtgattag	agctgtactg	atgtcattag	ccttgatggt	ggccaggact	ggtaggggcc	31740
tcagaggtca	tggagtccct	tcgtggagcg	ggtgctgagg	ctgtatcagg	cacagtgtctg	31800
gctgctttca	cctgggcccgt	ctcaccgaag	tgtccatgga	gcctgcgtag	ggtgggtatc	31860
tgtgtcgatt	ttacagatgc	agaaacaggc	tcagagaaac	cgagtgactt	ccctaaggctc	31920
acatacccag	ttagagcaga	gctgggcccag	gaagtgtctg	ctcaggctcc	tgaccaggctc	31980
tccttgcttt	gcactcttgc	caaaaccatg	atccagaact	gactttgagg	tccccggacc	32040
tcaggctcct	ccgaaatggc	ctcttgaggg	ctgctgagcc	acagcttagg	accacacctg	32100
agaggcaaat	gtgctttgag	ctgccaggcg	tcttgggggc	cctgccttgg	gcacgggggtt	32160
cagacaggcc	ccagatgtgt	ggggcgtctt	tctggacttg	agttttcttt	tctgtgtggt	32220
ggacacagtg	ctcacccctt	aaagcacctg	tgatgtgtgc	agcagcccaa	tccttgcttg	32280
tcgcctgttc	tgctagggaa	ggaaggaata	cttcaggatg	gcaggacaac	agaaagagggt	32340
ccagggtttta	gagcaagggc	aggtcaaact	tagaaaattc	tgggaatgagg	atgtgcattt	32400
cctcttctg	atctgtctaaa	agaagaggga	aggaggggct	gctggggggag	gagcccagag	32460
ccgagtttac	atccggatcc	cgcaaggcct	cccctgccc	gaggtccttg	tttgtgatgt	32520
gcttgtgtcc	atcctgggtt	ctgccgtgtc	cccaacatcc	ggccaagctt	aggtggatgt	32580
tccagcacac	actcaccctg	tctgtgcacc	tgtttttgtg	tccgtaagtg	ggtattttact	32640
caccttacga	gtgagccact	gtgggaattc	agggaggtgg	cgcagtgacc	acccttgagg	32700
ggatatgtgt	gtggcagggg	tcgagggtct	cgcccttccc	tgcttcctgc	gcgtggcttt	32760
ctccaggacg	gggagggctg	agctgaagag	gtggggacag	ttgcgtcccc	ccgccaccca	32820
ctgtcctgcg	gtgagagcag	actcactgag	cctgcccttc	tcccttgtgc	cttccagcta	32880

catctactgg	accgagtggg	gcggcaagcc	gaggatcgtg	cgggccttca	tggacgggac	32940
caactgcatg	acgctgggtg	acaaggtggg	ccgggccaac	gacctcacca	ttgactacgc	33000
tgaccagcgc	ctctactgga	ccgacctgga	caccaacatg	atcgagtcgt	ccaacatgct	33060
gggtgagggc	cgggctgggg	ccttctgggtc	atggagggcg	gggcagcccg	gcgttgggcca	33120
cctcccagcc	tcgcccagcg	tacctgtggg	cctgcaagtt	ccccaacctg	gcaggagctg	33180
tggccacacc	cacgactgcc	cagcagcctc	accctctgct	gtgggagttg	tccccgtcca	33240
cccctgggtg	cctttgctgc	agttatgtcg	ggagaggctc	tggtagacagc	tgtttcctgt	33300
gcacctgctg	ggcactaggt	cccagctaata	ccctgtgcca	ggactctaata	ttcacccata	33360
cacacatggt	ggttttcatt	gctggggaag	ctgaggcctg	agcacatgac	ttgccttagg	33420
tcacatagct	ggtgagttca	ggatccccc	gagataccag	ggccagcact	cgatccccac	33480
ccagccctga	accccacat	gtgctgggat	tgtgctggga	gtgtccacac	gcctgggacc	33540
ccagggctgg	tgctctcatc	tcctttttcc	agatcatgag	aatgaggctc	agggaagttt	33600
gaaaaaaacc	tatcccaagt	cacacagcaa	caggagcagg	atttgaaccc	agaaaagggg	33660
accgcacact	ctgttctgct	agagttagtta	gctgtcctgg	gtgatattggc	aggtgacagg	33720
ggcaactgtg	cttaacaaag	gaacccccat	ccccctgcc	aagttgggag	actagaagggt	33780
caggggcaga	agctctgaag	ggccagggtgc	agtggctgac	acctctaatac	ccagcacttt	33840
gtgaggccaa	ggcgggcaga	tgatttgagc	ccaggagttc	aagatcagcc	tgggtaattgt	33900
agtgagacgc	catctctaca	aaaaaattttt	ttaaaaatta	gctgggcatg	gtggttcatg	33960
cctgtagtcc	aagctacttg	ggaggctcag	gtgggaggat	tgcttgagcc	caggagggtg	34020
aggttggtgt	gagctgtgat	catgccactg	cactccagcc	tgggcaatag	agtgaagccg	34080
tctccaaaaa	aaaaaaaaaga	agaagaaaaa	aagctctga	ggctccaagt	ccccaggcac	34140
cccttggttt	gagggcagac	aaggaggag	agggtcacct	gggcagccct	gacttttgtc	34200
ccctggcaaa	gggaccttca	gtgaccttgg	ccctaggaga	gcctctgagc	acgtcagcca	34260
tgctgaaccg	ctcaggaagg	gcagcaagaa	tttggttct	gacctctgcc	tctcctactc	34320
gccatctgca	ctgggtgtgg	ttgtgccc	tttacagatg	aggaggctgg	ggcatcgacc	34380
agctgaatgc	cttgtcccag	gtactgcgta	ggcagagctg	gcagttgaac	cccgtgtcct	34440
ggttgctcgt	gggggtgggc	tgacccctga	cttgtgaggc	cagtagcaag	gtttgcacgt	34500
gacttcgtga	ccgtcaccca	gctctgcagc	acatcccgtg	accagctca	tccaggccgc	34560
atgcaaacct	gttgccaggc	gagaaaccag	tcaccgcaca	gctgtggttg	cctgaaatga	34620
ttaagctcat	taatcacccc	ggagttagga	cagactcaga	tgaaaaccag	caaaagccct	34680
ggaaactcat	gtgaccctgc	caatgagggc	ggccatgtgc	attgcagcct	ggccgtcact	34740
cctcggtacg	tgttttggtg	ttaaacgctc	cggatgttta	ctgagtgcct	gattaataac	34800
atggaaggcc	tggtctcatt	gctgtgggag	tgaaggatgc	acagccaggc	ctgacatgat	34860
gagaacaaga	acctggagtc	tcgctgcctg	gggtgtaatc	ctggccctgc	cacttagcaa	34920
ctgtgtgact	gtagccagggt	cacttaattt	tgctagatcc	tgctgcgct	tcagtggatc	34980
ttgtctggtt	tccaagggtg	ccaaacactt	taaggcattc	atgtggtcgc	taggctgcag	35040
ggttgaaccc	tggtcaccc	cgcagggcgc	cgtgtgctct	gtggcctggc	tgtgcctttg	35100
ctgacaccgt	gcccgtgtgt	gttcatgcag	gtcaggagcg	ggtcgtgatt	gccgacgatc	35160
tcccgcaccc	gttcggtctg	acgcagtaca	gcgattatat	ctactggaca	gactggaatc	35220
tgcacagcat	tgagcgggccc	gacaagacta	gcggccggaa	ccgcaccctc	atccaggggc	35280
acctggactt	cgtgatggac	atcctgggtg	tccactcctc	ccgccaggat	ggcctcaatg	35340
actgtatgca	caacaacggg	cagtgtgggc	agctgtgcct	tgccatcccc	ggcggccacc	35400
gctgaggctg	cgctcacac	tacaccctgg	acccagcag	ccgcaactgc	agccgtaagt	35460
gcctcatggt	ccccgcacc	tactcctc	gttagatcag	gctggttctg	ggagctgacg	35520
ctgaaaggag	cttctcatct	ggggttctctg	ggtgtacata	gatggttggg	taggttgtgc	35580
actgcacaag	ctgcatgatg	ctacctgggg	gtccagggtcc	aggctggatg	gacttgttgc	35640
ttcatcagga	catagataaa	tggccaaaac	tcctcagctg	gaaggtcctg	ggcaggatct	35700
ttgggtgtga	aaaccagtca	caggggaagg	gtgcttgctc	atactgccag	cacagtgtctg	35760
agtgttttcc	atagcgtctg	tttactcctc	aagcctggag	ggtggggagt	agcatggtcc	35820
catttcacgt	acaaggaacc	cgatgcacag	agaggtgtgg	caaccatcc	aaggccatac	35880
aactgggggtg	ggttgagccg	gggttgactg	tggcaggctg	gctcaagagt	ccctgctcct	35940
gaacccttgc	caggcagcct	ggcatcagct	cggggaattt	ttgccctgac	ccttggaagc	36000
aagtgggcct	ctttgttctc	atgtcagtga	tgagaagagt	gactttccta	tggccctct	36060
ggagtagcag	tgtttctctg	tggcgggctc	ttcccccatg	acatcagcag	cgagctggtt	36120

atgattccct	acgcagaact	tgatagttta	taaagctctt	tgtcatccag	gccccgttgg	36180
agtctcacgc	agacctgggc	gcaggcgggg	ctggtcttgc	ctgtcccagc	tgcattggatg	36240
gggaacttga	ggcttgcaaa	ggttaagggg	ctgttcgagg	cccacgttgg	caggagatgg	36300
gcctgggcca	gagtcctggga	cttcccatgc	ctgggctgtc	tttggctcctg	ttgctcacca	36360
tccctccctg	gggccatgac	cttagagagc	caaattggagg	tgcaggtaac	ccacggcaag	36420
gaggggttgc	catgactcag	agtccccgtc	ctgtggcccg	cagtacctgg	tgcaacgact	36480
tggatttcag	accagccact	gtagcccgtc	gacgggtgcg	tcgaagtgcc	acagcttctg	36540
aagccaggca	ggactcaggc	caggagactc	tgtagctgt	tgagaggag	aggccaacgg	36600
atgttctggg	tctgctagag	agctggttct	tcggatcctg	gtaccagtgc	actgagagga	36660
ggcccagctt	gattctgggg	ctgccttgtg	gtggcatgtg	ctgctcactg	acaccctcga	36720
ggagtgtctt	ctctcgggct	tgttgactgt	gccccgtttt	ccgcagttca	ctggtgcaca	36780
cataggcaca	tagcaaaccg	cacacacagt	cgtgggtatg	agtttacta	cattccacca	36840
ccagtgttca	ctaccattac	ctgccttccg	tcttaagtgt	tcattcattta	aaaataaatt	36900
tattgggctg	gacgcgggtg	ctcatgactg	ttatcccagc	actttgggag	gctgaggcgg	36960
gcagatcacc	tgaggtcagg	agttcaagac	cagcctggcc	aatatgggtga	aactccatct	37020
ctactaaaaa	tacaaaatta	gctgggcatg	gtggggcatg	cctataatcc	cagctactca	37080
ggaggctgag	gcaggagaat	ggcgtgaacc	cgagaggcag	agcttacagt	gagcccagat	37140
agcaccactg	cagtcacagc	tgggcaacag	tgcgagactc	catctcaaaa	aaaaaataaa	37200
taaataaaaa	aaaaataaat	ttatgatcta	tttcaaaaaa	aacacatgta	ctttgaaaca	37260
gcagagacac	atatgacacg	gagaatgaaa	ttccccatag	cgcaccccca	agagacagcc	37320
ctggtccccc	cgtctttccc	gtggacctcc	agcggggcag	atgctgagcc	gcctgttgct	37380
gagtggcatg	ctatcccgtc	ctccagctcc	tctgtggctt	acagacaccc	acctgcagcc	37440
ctgtctttgc	ctcctctagc	gcccaccacc	ttcttgcctg	tcagccagaa	atctgccatc	37500
agtcggatga	tcccggaaga	ccagcacagc	ccggatctca	tcctgcccct	gcatggactg	37560
aggaacgtca	aagccatcga	ctatgaccca	ctggacaagt	tcattctactg	ggtggatggg	37620
cgcagaaaca	tcaagcgagc	caaggacgac	gggacccagg	cagggtgccct	gtgggaaggg	37680
tgcgggggtg	gcttcccaag	gcgctcctct	tgctgggttc	caggctgctg	cccctgtcct	37740
tagcagaggg	aggaaacaga	ggatggctct	gggtgaatga	tgacttgggc	ttcgattatg	37800
tagtcacagg	gtatgacct	gagatgcgtg	gaaccccgag	actgtgatta	tatgtagaaa	37860
ctgggtttcc	ccgttggtta	agtagtcag	gtggggtcag	acccacaggg	acttttgtct	37920
tttcaagaaa	gaaaatgggc	gtgtgtcatg	caggggtagt	tggtactggg	taatccagggt	37980
ttatccttta	ttttgtggga	actgtacagt	catttctgct	acaatgctgt	atatgctcct	38040
ctgaaagaca	cctatgcaaa	atcgcacagt	aaaaatgaca	caactcatag	ggaaagcggg	38100
gccagggcac	agccctcaaa	atctccatca	atgacatgta	agaaaagaga	ggaaacctggg	38160
aaatagcaaa	gtgccttttg	cacattaaat	ggttagctat	atcccacaat	actgtgcatt	38220
cgtaaacgtt	aatgctgcaa	taaatacggc	acttcacctt	gggaagatct	ggagttggct	38280
tatgagtgtg	gaagggtgta	gcgcatgagt	ttttgtgaaa	cactggaagg	aggattgtgg	38340
gaaatcaaat	ggaaagtct	caccccaggc	gtggagaaga	gtgggtcatg	gccccagcag	38400
tgagcccagg	gaggtcagag	acggaggtgt	gtgtgtgggt	gtgacctgc	gcagttccct	38460
gccggctgta	gttttttgca	ttcgcttaat	gtttctcgtg	gaggaaattg	tgcattgagca	38520
aatgtgaaac	cgtgctgtgc	tcaaattgtc	ctaatacatc	attgcattgg	aacagattgg	38580
ctttnttttt	tttttttttt	tttttttttt	tttgaaatgg	agtctcactc	tgtcaccagc	38640
ctggagtgtg	gtggcatgat	cttggctcac	tgcaaccttt	gcctcctatg	ttcaagtgtg	38700
tttctgtcct	cagcctcctg	agtaactggg	attacagggc	atgagccacc	gcggccggcc	38760
agatttgcat	ttttgaaaca	actgctaggc	tgggcgcggg	ggctcacacc	tgtaatccca	38820
gcactgtggg	aggccgaggc	aggtggatca	cctgaggtca	ggggttcgag	accagcctgg	38880
ccaacatggt	gaaaccccg	ctctactgaa	tatacaaaaa	tcagctgggt	gtgggtggcg	38940
gtgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	aattgcttga	acccaggagg	39000
cagagggttg	ggtgagccga	gatcacacca	ttgcactcca	gcctgggcaa	caagagcaaa	39060
actccatctc	aaaaaataaa	aaatagaaaa	acaagtgtctg	tagcgggaagt	gagcactttg	39120
cggagtccag	cttgtgtggc	ctgttccaca	aatgatgtgc	tcacgggtgg	ctcaggccca	39180
cctggagtct	gcagcatggg	gcacaacagg	ttcattagtgt	tagaattcca	ggacaggcct	39240
ggctcctaag	cagccttctt	ttacaaaaaac	tgcagagccc	gcctgtatcg	tagcactttg	39300
ggaggccgaa	gtgggtggat	cacgaggtca	ggagtccaag	accagcctgg	ccaacatggt	39360

gaaaccccat	ctctactaaa	tatacgaaaa	ttagctgggt	gtggtggcac	gcgcctgtag	39420
tcccagctac	tccggaggct	gaggcagaat	tgcttgaacc	tgggagggtg	aggttgcagg	39480
gatctgagac	catgtcattg	cactccagcc	tgggcaacag	agcgagacgc	catctcaaaa	39540
aaaaaaaaacc	tacagagcca	cacggcctct	ttctccaccg	agtgttggtg	tgggagcttg	39600
tggtattgtg	gtgaaatctt	ggtactttct	tgaggcagag	agaggctgag	cgcctggaga	39660
gactttcaca	tgggtcgcca	tgtccgcctg	cggtttcgct	gttgtgctcc	ccatctgaag	39720
gctggtgccg	tccagacagg	ctggacgccc	ctttccacca	gatccttcct	cccgcagcag	39780
tttctagtta	cgttgtactg	tgaggctctg	gtccttgggt	gatggcaaaa	gtcagccgaa	39840
ttgaaattca	gagccatgcc	tggctccctg	gagcttctct	cctgggcagc	tgtgatcatt	39900
gcctctgctg	tgggtgtgggt	ggtggaaatg	gattcctttc	atcttgcttg	ctacagggtga	39960
ctgtcacgtg	gagtcctttg	gagagaggga	cgtgttaatt	gatggatgtg	gctcccatgc	40020
tgagaaaagct	cctgggcgta	cattgcctta	gagtttcatt	ggagctgcgt	tcttttatgg	40080
tgcttgctag	gcagaagtga	tgaagacttg	gaagaaaacc	cagaagggtt	tccacttaat	40140
ttggaaaatg	tgcttttccc	ctcctgtgtc	ttttgctaag	gtccagcctc	ctgcagcctc	40200
cccgtctgtg	ggactctggc	tttgattctt	tattaggagt	ccccctgtc	ccccaaaaga	40260
tggtgtctaa	attatcatcc	aattggccga	ggttttgttt	tctattaatt	gtttttattt	40320
tttattgtgg	taaatttata	taacataaaa	tttgccattt	taattgtttt	gttattgttg	40380
tttttgagac	agggctctac	cccagtgcct	aggctggagt	gcagtgggtg	gatcatggct	40440
cactgcagcc	tcagcctcca	gggctccagt	gacctctca	cctcagcctc	tctagtagcc	40500
gggactacag	gcatacacta	ccacatctgg	ctgatttttt	gtattttttt	tttattgtag	40560
agaccgcgta	tgttgcccag	gctggtctca	actcctggac	tcaagccatc	ctcccacctc	40620
accctcccaa	agtgtctgga	ttacaggcat	gagccacaac	accagcccat	tttaattttt	40680
tttttttttt	ttgagatgga	gtctcactct	atcgcccagg	ctggagtgc	gtggcgtgg	40740
atcaactcac	tgcaacctct	gcctcccagg	ttcaagcgac	tctcctgcct	cagcctcctc	40800
ccgagtagct	gggattacag	gtgcccatca	ctatgcctgg	ctaatttttg	tatttttttag	40860
cagagacggg	gtttcaccat	gttggccagg	ctggtcttga	actcctaacc	tggtgatccg	40920
cccgcctcgg	cctcccaaaa	tgctgagatt	acaggtgtga	gccaccgtgc	ccggcctttt	40980
tttgtttttg	agacagggtc	ttgccctgtc	accagactg	gagtgcattg	gtgggctctt	41040
ggctcactgc	agcctccgcc	tcccaggctc	aagttgtgca	cctccacacc	tggctaactg	41100
tattttatgt	agagacagat	ttcaccatgt	tgcccaggct	gggcttgaaa	tggactcaag	41160
cagtccaccc	acctcagcct	cccaaagtgc	tgagattaca	ggcgcgagcc	accgcaccca	41220
gcccatttta	cctattctgc	agttgacagt	tcagtggcat	tcagtccagt	cacgaggtaa	41280
ccatcactgc	cattcatctc	cagactactt	caccttctcg	gcagatgtcc	gaaactgtcc	41340
gcattgaaca	cactcctcat	ctcctctga	cagccacat	tctactttgt	atctctctct	41400
gccttctcta	ggtacctcat	gtaagtggaa	ttataccaat	atttgccctt	gtgtgactgg	41460
ctctttctta	gtgacatggg	gtcctcaagg	ttcatctgtg	ttatagcctg	gtgcagaatt	41520
tcttctctta	aagcctgaat	aataaccctg	tgtaaaggct	gggcgcggtg	gtcacacccc	41580
tctaattccca	gcattttggg	agtccgaggt	gggcaggtca	cctgaggtca	ggagtttgag	41640
accagcctgg	ccaacatagt	gaaaccctgg	ctctactaaa	agtacaaaat	tagctgggtg	41700
tgggtggcgg	cacctgtaat	cccagttact	caggaggctg	aggcaggaga	atcgcttgta	41760
ccggggaggc	agaggttgca	atgaaccaag	attgtgcctc	tgcagtccag	cctgggtaac	41820
agagtggagac	ttcctgtctc	aaaaaaaaaa	aaaatcatcg	gatggatgga	cggaccactt	41880
cttggtatatt	atccatccac	gggtgctagg	tttcttccac	ctttgggtgt	cgtgaataag	41940
gccactatga	acatttccct	ccgtggtgaa	ggttttgtac	tagtgaggaa	aaggcgtgtt	42000
tggtggtgtg	cataggattc	tggttaagaaa	gtttgcacta	accataagta	tttgtactac	42060
attaaaatga	aagctcaggg	gccggggcgg	gtggctcacg	cctgtaatcc	cagcactttg	42120
ggaggccagg	gcgggaggat	catgaggtca	ggagatcaag	accatcctgg	ccaacatggg	42180
gaaaccccg	ctctactaaa	aataccaaaa	aactagccag	gtgtgggtgg	gggcacctgt	42240
agtcccagct	acttgggagg	ctgaggcagg	agaatggcgt	gaaccgggga	ggcggagctt	42300
gcggtgagcc	gagatcgctt	cactgcactc	gagcctgggc	aacagagcaa	gactccgtct	42360
cacgcaaaac	tctgtctcac	gcaagactcc	gtctcaaaaa	aaaaaagagt	tcagggttta	42420
tgaaactggc	cagccgcgta	aagtttgctg	tggtgttttt	gtgcccggga	ggagtgtggc	42480
cagggtgtca	cgtcacacag	tacacgtttc	tcagatgggt	gttctccaga	ctgctgtccc	42540
aaagtctgtt	tttgcactct	gttcccacag	accaccctc	cacggtgagc	ctgatttttg	42600

ccagggtagc	tggaatcttg	cttgtctttc	agcccggcag	ctgtaccagt	ccaggggtcca	42660
cagctagtgg	cttttaggaa	ggaatttggt	cagttggctt	tgacacatgg	ccccctaggg	42720
tccacagctc	tgtagtgatg	tggatgttgt	tatctacaaa	gacacatgat	ccttcgtgtc	42780
cagatgaaag	tgatgatgtc	tttgcagctg	cccagcaagg	ctgtgtgtgt	gtgtgtgtgt	42840
gtgtgtgtgt	gtgtgtgtgg	tgtgtgtgtg	gtgtgtgtgt	gtgtatgggg	gagggaggca	42900
ccctttccat	ctgggggtgt	gtgtgtgtgg	gggtgtgtgtg	tgtgtgtgcg	cgtgtgtgtg	42960
gtgtgtgggt	tgtgtgtgtg	tatgggggag	gcaccctttc	catctgggtc	caagagactg	43020
ggcctgggga	agacgcttct	ttttatctac	ttagagactt	tgttttattt	gtattttttt	43080
gagacagggt	ctcactctgt	cacccaggct	ggggtatggg	gatatgagca	tagctcactg	43140
cagcctcggc	ctcccaggct	gaagcgatcc	tcccacctca	gccttctgaa	tagctggggac	43200
tgtaggcgtg	cgtcaccata	ctgagctatt	gttttttttg	tttgggttgg	tttaattttt	43260
ttgatacaga	tggagtcttg	ctatgttgcc	cagactagtc	tcaaactcct	gaactcaagt	43320
gattctccca	cctcagtttc	ccgacattct	gggatcacag	gtgtgagcca	ctgctgtctc	43380
cctgttttat	taactgctga	aagacctaga	taaagaaagt	ctgaaaagac	ttactatcag	43440
agcaccatcc	taagatgatt	ccctctgact	caatggagag	ggaggggagc	ttttccttca	43500
ggcctgggtg	gcaggagccc	aggtgctcca	ggccccattt	gccccaggcc	aaatcactcg	43560
ggaacttgga	tgcagctgtc	tttcagggtg	acccaaagga	accagatccc	cgcaggcagt	43620
aggcttctgg	gctgtcctct	cctcctacgt	cagctcagta	agagcccttc	gaagggatgc	43680
tgtgtcggag	gccccaaaag	cccaggctca	tccctgagat	gcacagggtg	ggctgggctt	43740
aggcagcgct	cgagcatctc	ctggacgggtg	acccagagag	gtgtggagac	ggagagctct	43800
tgagagtcac	tgagagacgt	ggctgcccctg	ccttcccagg	aggggctctg	agtcattccc	43860
cacactcacc	tgccccctacc	caccctcacc	tggccccccag	cctcacctac	ccccacatct	43920
gtaccgatcc	ctttaccogc	accttccccta	cccaccctca	cctccccctgt	accttcacct	43980
ccccactca	cccgcctctg	caccctcacc	tgtccccccac	cttcacctaa	ccccaccct	44040
cacctgccct	cccctcacct	ggcctccttc	cgttggggaa	ggggttgtaa	ggggcggccc	44100
ccaaactgtc	tgtcctgggtg	ccctgcagag	aaaacagtac	gtgagggccg	cagtccaaaa	44160
gcttgagtcc	tggaagggtg	aggagacagg	gatgtgttgg	gaagggcccc	atggctcttg	44220
atcccttctc	gactgtcaat	ggggccttca	tgggagcgcc	agtctagtga	tgcacagctg	44280
gggtgccggc	gggtggctga	ggaggcctaa	agtccgaggc	ggcaagagct	cttcagagg	44340
ctgttgtcct	aatcgctctg	gcatactcag	gcgggcacgt	agttaggagc	tgattggaga	44400
ggagagaccc	ccacaccaat	actgggattt	gactttcagg	ctaaacttga	gaagtgtggc	44460
ctctgctgtc	ctgccagagc	tctccagcca	gtgcccaggg	ctctccagcc	agtggccggg	44520
gggtccacc	agtggccggg	gggtcccgcc	agtggcaggg	gtctccgcca	gtgcccaggg	44580
gtctccgcca	gtgctcagga	gtcttgggtt	ctttgtctta	cagccctttg	ttttgacctc	44640
tctgagccaa	ggccaaaacc	cagacaggca	gccccacgac	ctcagcatcg	acatctacag	44700
ccggacactg	ttctggacgt	gcgaggccac	caataccatc	aacgtccaca	ggctgagcgg	44760
ggaagccatg	ggggtggtgc	tgcgtgggga	ccgcgacaag	cccaggggcca	tcgtcgtcaa	44820
cgcggagcga	gggtaggagg	ccaacgggtg	gggtgggggtg	ctgcccgtcc	aggcgtgccc	44880
gccgtgtctt	ctgccgaatg	ccagcctctc	acaggctggg	gagactttcc	accctgggga	44940
tccaatgggt	ggctttccag	ggccccaaaa	gcaaacacag	gctctttcac	agccccctca	45000
ggaaagcaga	aagccccaa	ggctggaaag	gaagggggag	ctctgctgag	agggttacaag	45060
gcagcgctgg	ccgacgggag	ttgcagttga	taggttttgt	atcatccttg	ttaaacttga	45120
accctgtgca	gaaatccctt	ccacggcatg	ggggctgcct	gttgactcgc	tcctgttcca	45180
ccacaggggag	ctcctgggct	tcttctctcc	agaggccccc	gacgtcccca	cctgttgggtc	45240
gtcagagctt	ctgggttgggtg	ggaaggcacc	caggaccttg	aggctctccag	agagaaaagc	45300
cagggaaaaga	gggagaccga	aacctatgtg	acatgaaact	caggctccaa	actgagcacg	45360
ggaacgtttg	gggacaggag	cgcgatggcc	ttcctcagat	agctgggggg	ctggcatgaa	45420
gacgggagct	acagccagca	caggctcctg	gccgggagcc	cagagattga	gccctgactc	45480
tgtcacttac	tggccacgtg	accttgggctg	gggtggcatag	cctcttggag	actcagtttc	45540
ctcattggtg	ggagtgcagg	ccacagtggg	gcggcctctg	cagcacacgg	ggggctcggt	45600
gggcggaagc	ccgggtcta	taaggcggct	gtgcaggagc	cagccgagct	gggtctcccaa	45660
cagccagggc	tccgggtcc	ttagcagctg	tggggggcct	gcacctgttt	cccatggctg	45720
ctgtcagaaa	ttaccagaag	ccagggtggct	gagagtaatg	gacacttggt	ctctcacagt	45780
tcctgagggc	tgaagccga	gatcgagggtg	tgggcagggc	cctgcgcctt	ctgaaggctc	45840

tgagggaacc	tttgggcttc	tggtggctcc	aggcaccctt	tgacttgtgg	tcctgtcact	45900
ccagtctctc	tgtctggctg	cacatggcgt	ggcctcttct	gtaccattga	aggacacttc	45960
agttggattt	agggcctacc	ctcaccattt	gtggctcgat	cttgatcctt	catgacattt	46020
gtaaagaccc	tgtctccaaa	taagctcaca	ttctgaggtt	ctggggtgag	cgggaatttg	46080
gagagcattg	ttcaactagt	atagaatgtg	acctgtcagc	ctcgggcagc	cctgagaggc	46140
aggggctttc	cacagcccag	ctgggtgccc	tgggctccgt	gctgtccgag	gagacgccat	46200
ccccacaccc	gtccttcacc	cgccaccctc	ccgcaggtag	ctgtacttca	ccaacatgca	46260
ggaccgggca	gccaagatcg	aacgcgcagc	cctggacggc	accgagcgcg	aggtectctt	46320
caccaccggc	ctcatccgcc	ctgtggccct	ggtggtggac	aacacactgg	gcaagctggt	46380
ctgggtggac	goggacactga	agcgcattga	gagctgtgac	ctgtcaggta	cgcgccccgg	46440
ggcctgccc	aaccgcagac	acccggcctt	cattgtcagt	aatggcagca	gctgccacat	46500
tgtccgagac	ctgccgtgag	cccagtgccg	cgccaggggc	tttgtgtgta	gcgtgttttg	46560
tcctcacact	gacagctgta	ggctgggggt	ctgagtggag	cccacagggc	agaggcagaa	46620
aatgagtctc	agagagggtg	agcagctgc	ttggggcccc	acagcaggag	atggagcagg	46680
actgcagcct	agcctctgcc	cccagcacct	gcgcaagaag	ctgctctgct	ctggactgtg	46740
ttaggctgcg	agggctggag	agaaatgaga	gttgggtgctt	agagaggggg	cgcagggtccc	46800
catggctttt	cctcttatga	tgaggtagat	gggtgaaggg	aggggcatag	cttgacagggg	46860
ccagtgaccg	aggcccggcg	ttggaactga	tggccttcac	cccagagcca	gcccagggtg	46920
gagcagggtt	ttccgagggc	ttgtcttggg	tcggcctgct	tccagggact	ctgctgcagc	46980
tcccacccct	gtccaaagca	tggaaatccc	caggctccct	ggcagtcctg	tcaacctctg	47040
tcctcccaag	ctgagtgtgg	ggcaagtctt	ggaggtcagc	actgctcagg	ggggcccacg	47100
ggctgcttgc	agggggccaac	ggcctgaccc	tggaggacgc	caacatcgtg	cagcctctg	47160
gcctgacat	ccttggcaag	catctctact	ggatcgaccg	ccagcagcag	atgatcgagc	47220
gtgtggagaa	gaccaccggg	gacaagcgga	ctcgcaccca	gggcccgtgtc	gcccacctca	47280
ctggcatcca	tgcagtggag	gaagtcagcc	tggaggagtt	ctgtacgtgg	gggctggcag	47340
tgggggtggg	aggggtggcct	ctaaacccga	cccctggagg	aggctggagg	ccagtgcag	47400
atcctgtgtg	gcctcagcca	ggcgggtggtc	tctgccagat	gccaaactgtt	gcccgcagg	47460
gttcagcgac	atgtccgaat	gtcccagggc	ctctgaggtt	gttttctttt	gcccgcagaa	47520
aaatcaccac	gaacagcggt	ttaagacaac	accaactctt	tttttttttt	tttttttttg	47580
gtcaggatct	tgtctctgtt	cccaggctgg	ggtgcccttg	tgcaaacaca	gttcactgca	47640
gcctcgacct	ctgggcttaa	ttaagtgaac	accttgccct	agcctcccag	gtagctggga	47700
ctacagggtg	gcaccaccac	acctggctaa	tttttttttg	tagagacggg	gtttccccat	47760
gttgcccagg	ctgggtctgca	actcctgggc	acaagctatc	tgccctgctgt	ggcctcccaa	47820
agtgttagga	ttataggtgt	gagccactgg	cctgacaaca	cccacggatt	gtctctcagt	47880
tctgtaaggc	aaagtccagg	cacagcgtgg	ctcacctggg	ttctctgctc	agggctctcac	47940
ggggccagaa	tcaaggtgtc	aggaacgctg	ggccctcagc	ggaggctctg	tggagaaatt	48000
agcttccttg	ctcactcagc	aggtagcagt	tgtgggatcg	aggttctgtt	ttctctctgg	48060
ttattggctg	gggaccactc	tcagctccta	gaggccaccc	caggctcctg	ccccgtggcc	48120
ctctctgcct	cagcagtggg	ggctccctgc	gtcagtcctt	cccgcacctt	gagtctctct	48180
gatttgcttc	taaagggccc	tgtgattcgg	ctcagccacc	tttagattag	gttagcctcc	48240
cctttgatag	actccaagtc	ggctgattaa	taaccttact	cacatctgca	gaatcccttc	48300
tgccacataa	ggctcatgac	ccgtgctggg	gactgggggtg	ggaaattacg	gggtcattta	48360
ggattctgcc	tgccactgcc	ttgctgtgtc	ccagggtctg	ggggaggggc	ctccacagct	48420
gggaccacag	tccttcctcc	cctccatggt	aaccatctga	ggattacttg	agaccagcct	48480
gggcaacatg	gtgagaaccc	atccctacaa	aaaatacaaa	caaaaaggga	ccaggctggg	48540
cttgggtggc	catgcctata	atcccagcac	tttgggagac	caagggtggg	tgatcacttg	48600
aggttggggg	ttcgagacca	gcctgcccac	catagtgaac	tcccgtctct	actaaaaata	48660
caaaaattag	ctgggtgtgg	tggcaggcgc	ctgtattccc	agctactggg	gaggctgagg	48720
tgggagaatt	acttgaacct	gggaggcgga	agttgcagtg	agccaaaatt	acgccactgc	48780
actccagcct	aggcaataga	gtgagactcc	gtctcaaaaa	aaaaaaagg	ccagggtgtg	48840
tagtgacaaa	gagaccctat	ccccaaaaaa	ccgaacactg	aatccttgag	actgagtaag	48900
gacactgtga	aatttttctg	ggtggggcag	ggaacagagc	gtcttctgtc	atttcttcca	48960
cctgggtgtg	gtcagctctc	cctccaagct	gcctcctctt	cttctcattg	tccgggtgtt	49020
ggacacattt	ggttaactgg	atagaataac	gcgagttccc	agggacttgg	tccatttgct	49080

at t t t t a t t t t	at t t t t a t t t t	at t t t t a t t t t	at t t a t t t a t	t t a t t t a t t t	at t t a t t t a t	49140
t g a g a t g g a g	t t t c g t t t t t	g t c g c c c a g g	c t g g a g t g c a	g t g g c g c g a t	c t c g g t t c a c	49200
t g c a a c c t c t	g c c t c c c a g g	t t c a a g t g a t	t c t c c t a c c t	c a g c c t t c c a	a g t a a c t g g g	49260
a t t a c a g g c a	c c c a c c a c c a	t a c c a g g c t a	a t t t t t t t g t	a t t t t t a g t a	g a g a c g g g t t	49320
t t c g c c a t t t	t g c c c a g g c t	g g t c t t c a a c	t c c t a g c c t c	a g g t g a t c c a	c g c a c c t c g g	49380
c c t c c c a a a g	t g c t g g g a t t	a c a g g c a t g a	g c c a c c a c g c	c t g g c a c c a t	t t g c t a t t t t	49440
a a t t c c c a t g	t g t a t t a g t g	t c c c a c g g c t	g c t g t a a c a a	a t g a c c a c a a	a c t g g a t g g c	49500
t t a a a g c a a c	a g a a a t g g a t	t c c c c c a a t g	t g c t g g a g a c	c a g a a g c c t g	c g a c c a a a c t	49560
g t t g g g a g g g	c t g t g c t t c c	t c t g g g g g c t	c c a g g g a g g a	t c t a t t t g t t	g g c c c t t c c a	49620
g t g c t g t g g g	t g c c a g c g t t	c c a c a c t t g t	g g a t g c g c c g	c c t c a a c c t c	t g c c c a t c t t	49680
c a t g t g t c c a	t c t c c t t t g t	g t c t g c g t c t	t t a c c t c t t c	t t c t t g t c t g	t g t t g c c t c t	49740
t a t a a g g a c g	t t t g t c a t t g	g g t t t a g g g c	c c a c c c a a a t	c a t c c g a g a t	g a c c t c g t c t	49800
t g a g a t c c t t	a a c c t g c a a a	g a c c c t t t t t	c c a a a a a a a g	g t t a t g c t c a	c a g a t t c t a g	49860
g c c t t a a g a c	a t g g g t g t a t	c t t t c t g g g g	g g c a c t a t c c	a a c c c c t t a t	a c a a t g a a a g	49920
a c g g g a a g a g	g g c c a g g t g t	g g t a g t t c a c	g c c t g t a a t c	t c a g c a c t t t	a g g a a g c t g a	49980
a g c g g g a g g a	t c a c t t g a g c	c c a g g a g t t t	a c a a g t a g c t	a g g c a a c a t g	a t g a g a c c c c	50040
a t t t c t a c a a	a a a g t a a a a a	a a a a a a a a a a	a a a a a a a a a g	c c a g g t g t g g	t g g c t c a c a c	50100
c t g t a a t c c c	a g c a c t t t t g g	g a g g c t g a g g	c a g g c a g a t c	a c g a g g t c a g	g a g a t t g a g a	50160
c c a t c c t g g c	t a a c a c g g t g	a a a c c c c g t c	t c t a c t a a a a	a t a c a a a a a a	t t a t g g c c g g	50220
g c g c a g t g g c	t c c c g c c t g t	a a t c c c a g c a	a c t t g g g a g g	c c g a g g t g g g	t g a a t t a c a a	50280
g g t c a a g a g a	t c g a g a c c a t	c t t g g c t a a c	c g g t g a a a c	c c c a t c a a g a	t c a c a a g g t c	50340
a a g a g a t g g a	g a c c a t c c t g	g c t a a c a c g g	t g a a a c c c c g	t c t c t a c t a a	a a a t a c a a a a	50400
a a t t a g c c g g	g c a t g g t a g c	g g g c g c c t g t	a g t c c c a g c t	g c t c g g g a g g	c t g a g g c a g g	50460
a g a a t g g c g t	g a a c c c g g g a	g g c g g a g c t t	g c g g t g a g c c	g a g a t c g c t c	c a t g c c a t t g	50520
c a c t c c a g c c	t g g g t g a c a g	a g t g a g a c t c	c g t c t c a a a a	a a a a a a a a a a	a a a g a a a a t t	50580
a g c c a g g c a c	a g t g g c a g g t	g c c t a t t g t c	c c a g c t a c t t	g g g a g g c t a a	g g c a g g a g a a	50640
t g g c a t g a a c	c c g g g a g g t g	g a g t t t g c a g	t g a g c c g a g a	t c a t g c c a c t	g c g c t c c a g c	50700
c t g g g c g a t a	g a g c a a g a c t	c t g t c t c a a a	a a a a a a a g c c	a g g c a t g g t g	g t g c a t g c c t	50760
g t a g t c c c a g	c t a c t c a a g a	g g c t g a g g c a	g g a g g g t t g t	t c g a c c c a c g	g a g a t c a a g g	50820
c t a c a g t g a g	c c a t g a t c g c	a c c a c t g c c c	t c c a g c c t g g	g t g a c a g a g t	g t g a c c c t g t	50880
c t c a a a g t a a	g t a a a t a g g a	g g a g a g a c a a	g t g g g c a g t t	c a g a c t g a t g	g t a t g g g c a c	50940
a g t a g a g a c t	g g t g c a g a c a	g g c t g g c c t g	t g a t g t c a a g	c a a c t t c t g t	a a c t g t t t c c	51000
g g c a t c c a t t	t g t g t g t c a a	t t t c c g t g t c	a g t a g g a a g a	c t c t g t a g g c	t g c c a a g a g g	51060
a a t a a g t g g g	a g g a t c c t c c	c a g a g a g g c c	g g g c c t g c a g	g a g g g c c a g t	t c t c a t g a g t	51120
t c t t a t t t g g	c c c c t a c c c t	c c a g g c t g t g	g t t c t g a g g t	g g g a g a c a g a	g c c t g a c c t c	51180
t g t t t g t c t t	g t t t t g t c t t	t g c a g c a g c c	c a c c c a t g t g	c c c g t g a c a a	t g g t g g c t g c	51240
t c c c a c a t c t	g t a t t g c c a a	g g g t g a t g g g	a c a c c a c g g t	g c t c a t g c c c	a g t c c a c c t c	51300
g t g c t c c t g c	a g a a c c t g c t	g a c c t g t g g a	g g t a g g t g t g	a c c t a g g t g c	t c c t t t g g g g	51360
t g a t g g a c a g	g t a c c t g a t t	c t c t g c c t g c	t a g g t g t g t c	c c t g g c a t c c	t t t t a a a a t c	51420
a c a g t c c c t g	t g g c a t c c a g	t t t c c a a a g c	t g a t t g t c t c	t t c c t t t g c c	c t c c t t t c t t	51480
t t c t a c t a t g	t g c a t t c g g t	g c t a t g a a t t	t t c c t c t a a g	t a c t g c g t t t	c c t g c a t c t c	51540
a c a a a t t t t g	t t a c a t t t t c	a t t t t c a g g t	a g t t t g a a t a	t t t t t a c a c t	t c t c c t g a g a	51600
t g a c a t c t t t	g g c t c a t g t g	t t a t t t a g a a	g t g t t g c t t a	g t t t c t a a a g	a g t t g g g g c t	51660
t t t c c a g c t g	t c t c t c t g c a	a c t g a t t t c t	a a t t t a a t t c	t a c t g t a g t c	t g a g a g c t t a	51720
t t t t a t a t g a	t t t c t g t t a t	t t t a a a t g t g	t t g g g t g t g g	t g t t t t t g t t	g t t a t t g t t t	51780
t t g t g t c t t t	t t g t t t t g t t	t t g c t t c g t t	t g t t t t g t t t	t t g a g a c a g t	g t c t t g c t c t	51840
g t c a c t c a g g	c t g g a g t g c a	a t g g c g c g a t	c t c a g c t c a c	c g c a a c c t c t	g c c t c c c g g g	51900
t t c a a g t g a t	c c t c t t g c c t	c a g c c t c c t g	a g t a g c t g g g	a t t a c a g g t g	c a c g c c a c c a	51960
t a c c c a g c t a	a t t t t t g t a t	t t t t a g t a g a	g a c g g g g t t t	c a c c a t g t t g	g t c a g g c t g g	52020
t c t c g a a c t c	c t g a c c t c g t	g a t c c g c c c a	c c t c g g c c t c	c c a a a g t g c t	g g g a t t a t a g	52080
g c g t g a g c c a	c t g t g c c t g g	c c a t t a g g t g	t g t t t t a t c a	c c c a g c a t c a	t g c a g t t t a t	52140
c t t g g t g a a t	g t t c t g t g t a	c t c t t g a a a a	g a a t g t g g a t	t c t g c t g t t g	t t g g g t g g a g	52200
t g t t c c a g a a	a c a t c a a t t a	g a t c c a g t t g	g t t a a t a g t g	c t c a t c a g g t	t g t c t c t a t c	52260
c t t c c t t c c t	g a c t g c c t g c	t t g a g c t g t c	a g t t a t t g a c	a g g g g t g t g g	a g t c t c c a a c	52320

tctaattggtg	gattttgttta	tttctcctag	tagttctatc	tttttctctc	cttctaccct	52380
tgatcctctt	ctccccctag	ggcttcctgg	tgttggtggt	gggagagtgg	ggtagtgaag	52440
aacctggact	ttagggccaa	agaggccagg	gttcaaattc	tggctctgtc	acttcccagt	52500
tgagtgaacc	tggctggtgc	ctgaatctct	gtgagcctcc	acttctctct	ctgtgaaatt	52560
gagagcactt	acctggcagg	ctgtcatggg	catcaagtaa	cagggcactc	cacctggacc	52620
ctgacacgtg	atgcacagga	atgccagctg	ctatgccatg	ggtgtggcag	tagtaataaa	52680
gtgaccatct	gtatcctcac	cacagtgaag	cctgtccagg	gctttctctc	ctatgcccc	52740
atgcctccag	gtggccttgg	atcctgttgg	ttctgtgtct	tgctcagcga	cctttctccc	52800
gtgggagttc	ctgggggttc	agcttcatcc	tacagacagc	agcacacact	ggctgtgcac	52860
cctttttttt	tttttttttt	ttttttttga	gatggagtct	cgcttttttc	gcgcaggctg	52920
aagtgcagtg	gtgtgatctt	ggctcactgc	aacctctacc	tcctgggttc	aagtgatttt	52980
cctgcctcac	cctcccaagt	agctgggatt	acaggctccc	accaccacgc	ccggctaatt	53040
tttgtatttt	cagtagagat	ggtgtttcac	catgttggcc	aggatggtct	tgaactcctg	53100
acctcagggt	atccgcccac	ctcagcctcc	caaagtgcag	ggattacagg	cgtgagccac	53160
cacacccgga	gtgccggttg	tttttagcag	tttgtcttgt	tcctggagag	actggctcct	53220
gcccaggagc	tggggagta	gggcccgggg	gtgctgcctc	acacctcgag	tttggccgta	53280
agcagagggg	acatttttgt	actgtcccc	tcctgagctt	cccagcagct	tttctccaag	53340
ttacagccca	aaagctcagg	tggatttgca	acccaacggg	gtctgtgcac	ctcccactga	53400
tgcccgaact	gccctggcca	agaaacgggg	ccgtcagaac	gctgcactaa	ctgcagcctt	53460
gggcctccat	gccagaggcc	atgcccttcc	atccaccacc	ccctggcctg	ggccctggcc	53520
ctcctggctc	gggaactcca	ggccccttcc	tcacggatcg	agagacgtgt	atttaccgca	53580
caggtgcttg	tcattctctt	gtggcctctt	ctccagggag	atcacagaag	gacagggcct	53640
cactgaggtc	tccgacatgg	accctttgat	agtggcagga	gccaggctgg	gcaagaggcg	53700
gccacagtca	cctcagcagt	gccatcacca	ccgccattca	gcccttcctt	gagccggggc	53760
cgcccctggc	tctggcccca	gtgtcccagt	tacagctcac	aggagcttgt	ggtgcccagc	53820
ggctgcttct	gattgagagt	cgaggctcga	ggctttggga	ggctgagagg	ctgctcgggt	53880
tcacaactgc	tgaggagagc	ttgggctcca	tctcaggtct	gccccatgtc	gccctcaacc	53940
tccagccacc	ggtcctccgt	gtcccccatg	gccaggcacg	gcttgcagac	atctgtcgtt	54000
ggctcctctc	agccgtcgtg	ggctgaccct	ggcacgtcct	cctgtggctg	agcccagtgg	54060
ggacagctgc	ttccttttat	taccctagaa	ctctcgtctt	tgatcaggcc	ccctccccta	54120
tgccacacag	tccctgtcac	tcgggtgagc	ccagtagtca	tggggaaggc	ctgcgggttc	54180
caaacatcca	aaggcttgcg	tgcagcatga	cagcttgaaa	ccgatgtttt	ttaccttgat	54240
cagatttcag	ccttggcgggg	gcttttgetca	gctttcagtg	aggcctgggc	cgatttccca	54300
gcatccctct	ctgaggccag	cctctgtttc	ctgtgatttt	ctgcacaaag	tgggagggag	54360
gagtcttagg	aaatgggggg	ccacctcgaa	acctaggcct	cctctggctt	ctctgtgcca	54420
gtgccccac	gcttttgtgc	tgtgtcccca	gcccatggga	ctgtgttatt	ccctgagtgc	54480
tgccgcagtc	ccagcccga	ctgaggacgt	ggagcccga	ggggcaggat	ggcctccatg	54540
gtcacacgta	ggaagtggcc	tccacctctc	gatgatcctc	ttccccctct	cctttcagcg	54600
ccttccccgg	gggtgtcatc	agccctcctg	cctgtgcttt	gtcccgtctt	ctgcaggcgc	54660
atgggacgtg	ctgacaggtc	ctctgcgggg	ttcctgcctt	gctatgcgca	cgctggtcac	54720
cacagaggcc	tggcccttct	tctgtagcag	tcccacacc	gcaacagggt	tggctgctga	54780
ccacctgctt	tctgccccct	tggctcctgag	gagggcgag	tgggcactca	ggcgtggctg	54840
agcagatgtg	tgttgccggg	aggaggaagg	actgctccag	tcagggtga	atttcccacc	54900
cggagcattt	ctgctgtatt	tgggtgtagc	cctgctgctt	aaagctctga	ttcccagttg	54960
gcaccctttc	ccttctgcat	tgaaaaaacat	acggatgcat	gtcttcttgc	agtgaatgtg	55020
tattctccca	gcctctcttc	tgggttgggg	ctggaggtgg	agcggcacac	aggagccgca	55080
gcgatggagg	atgtgcccgt	gcagcacccc	gtacagcagg	gatgccaaac	ccgcgctgag	55140
tccctctcaa	cttctgcttt	gaagcccagt	cacgccattg	cctgggtttt	gctgggcggg	55200
gctgcatgtg	atgttctcct	ctgtccctcc	cccagagccg	cccacctgct	ccccggacca	55260
gtttgcatgt	gccacagggg	agatcgactg	tatccccggg	gcctggcgct	gtgacggctt	55320
tcccagagtgc	gatgaccaga	gcgacgagga	gggctgcccc	gtgtgctccg	ccgcccagtt	55380
cccctgcgcg	cggggctcagt	gtgtggacct	gcgcctgcgc	tgcgacggcg	aggcagactg	55440
tcaggaccgc	tcagacgagg	tggactgtga	cggtgaggcc	ctccccgtca	aggctctgcc	55500
aagaccctgg	ccctgcccct	cgggatacga	gcttggggct	gcctccggcc	tcacaggagt	55560

aggggctctg	aaaacctttg	cttgcagggg	gattgccaag	tctgtctttt	aggcccaaca	55620
aggaaaaactc	tgcagttcca	cccacctctgt	cccaccaggt	agtgtggctt	gaaggcagac	55680
tgtgagggtc	tatctcacct	tcctgcatta	ggtcaggagt	ttcacagaaa	cctgaggcac	55740
attcaggggt	gggctgcaga	gggccatggc	tcacaccctg	gaaaatccgc	ccccaaaaga	55800
cagtgtgtgc	tccactgacc	agtctgtggg	atagtgttta	agcctgagtg	gtttctatca	55860
acatgtagaa	tcaggaggta	taaagagatt	tgtcagggca	tcctggggcc	tctctgacca	55920
gcaggatctt	ccttttagatc	ttgacagtga	aacacatctc	ttctgtgccc	cctgtgagtt	55980
ttctttcatt	cattcattca	ttcattcatt	cattcattca	ttcgagacag	agtcttgctc	56040
tgtcaccag	gctggagtgc	cctgggtgtaa	tctcggtca	ctgcaacctc	tgccctccagg	56100
gttcaatcga	ttctcctgcc	tcagcctccc	gagtagctgg	gatgacaggt	gcgcaccacc	56160
atgectggct	aatttttgta	tttttagtag	agacagggtt	tcaccatgtt	ggccaggctg	56220
gtctcgaact	cctgacctca	ggatgatccgc	ccgcctcagc	ctcccaaagt	gctgggatta	56280
caggcatgag	ccaccgcgcc	cggcctgagt	tttcctttta	tgaaggacct	gcttggttgg	56340
ttgcctgcca	catgttgta	gcacatggg	cccaggactg	ctgaggagct	gttgatgccc	56400
tcgtctccc	agagccaccg	gctctgttag	ataattcaca	tgcagtctgg	ccactgtcct	56460
acgtcctcat	tcacaaaagag	cagacatttc	gtagaagatg	agggcctggg	agtaacctcc	56520
ctgcatgttt	ttctataaag	gcatagtggt	taagtccttc	cagctcattg	accattggag	56580
aattttatgg	aggctgtaga	ctaggggctg	gtaaaactaag	ggcccagggg	ccaaatccag	56640
cctgccacct	acttttgtaa	ataaagtttt	cttgggtgcac	agccatgccc	attcattcat	56700
ttgcacaatg	tctgtggctg	ctttcatgcc	aaaagcagga	gaactgagtg	gttatgctgg	56760
agacctacgg	ccttcaaagc	cccagacctc	acgtctggcc	cttgacagac	agagcttccc	56820
cagccctgct	gcgcacccctg	gccagcatg	tgtgtgtgtg	gtgatttcag	cttgaggag	56880
ccgtgggttag	gaattgtccc	tgtgttggtc	cattttgcat	tgctatgaag	gagcacctga	56940
ggccgggtag	attatgaagg	aaagaggtct	gtctggctca	tggttctgta	ggcagcacca	57000
gtatggcacc	cgcactctgct	cagcttctag	tgagggtctca	ggaagctttg	actcatgggtg	57060
gaagtcgaag	cgggagcagg	tgcatacat	ggtagagagag	ggagcaacgg	agagagagag	57120
agagagagag	agagcgcctc	tccctcttgc	cctcaccttg	agaggagatg	ccaggctcct	57180
ttaagtaacc	agctcccatg	tgaactcaca	gtgagagccc	atttgctact	gcggagagggg	57240
caccaggcat	ctgctcccat	gacccaaaca	ctgcccacca	ggccctacct	ccaaccttgg	57300
ggtcatattt	tattctgttc	tatgctatgc	tatgctatgc	catgccatgc	catgccatgc	57360
tattcctatt	ctattatttg	agacagaatc	tcgtctgtgt	gcccaggctg	gagtgcagtg	57420
gcatgatctt	ggctcactgc	aacctccacc	tcacagggtt	aagcgattct	cctgcctcag	57480
cctcccaggt	agctgggatt	acaggcacac	accaccacac	ccgggtaatt	tttgtatttt	57540
caatagagat	ggggtttcac	catgttggcc	aggctgggtct	caaactcctg	gcctcaagtg	57600
atccacttac	ctcggcctcc	caaagtgcga	tgattacaga	tgtgagtcac	tgccgcccagt	57660
gagggtcaca	tttccgttga	gatttggagg	ggcagacgtt	ggagccatct	gagccccctc	57720
gtcccgtctt	agcttctcct	cccgtgtgcc	ccgcgggtgct	ggtaggcaggc	ccttacgcgcg	57780
gttctggctg	cacgctctgt	tccagaagct	ttcttccctg	cctgggttacc	agaaaatcat	57840
cccatccatt	acaaggacag	ggtcccctta	tctcccattc	ccagggcagg	acaccggggg	57900
cagggcagggt	ggggaaactga	gcaagtcttc	tgggggcagg	cgtggctatg	gctccctctg	57960
ggtgggcgtc	tggggagggg	tggaggcagc	cgtcagcgcc	ctggcttctg	cttccctccct	58020
ggccagagac	tgtggccttg	tgtgtctccc	gtgtgggctg	cctgcacctc	cagtgggttg	58080
tgtccctcc	cctccctcc	cctcaagctc	tgtgtgagc	caactgcctc	cacagcccc	58140
actctcggga	ggcgaggctc	ctcgtggcca	ttcctgtcct	tggcaccac	ccccccacca	58200
acctggtaga	gccttgggag	gggtctgtta	ctccttgcat	ggcgtagacc	tccccacagt	58260
aggcacctga	cacatacctc	ctggggggca	ggcaggaggt	gcgttgaggt	ctcagccctg	58320
gcagtccctc	ccttgcgtgg	cataggcctc	gccacagggt	catcgagggt	gggtggagac	58380
tgtactagac	cactccccgc	tggctcctaga	aagggtccca	tctgtctgct	ctctgttttg	58440
agtccagacc	ttgggtgctg	tgccctgcat	ggtgggctgg	ggggcaccct	ccagcctctc	58500
tgagtgcattg	gcctctcctt	gcagccatct	gcctgcccac	ccagttccgg	tgtgcgagcg	58560
gccagtgtgt	cctcatcaaa	cagcagtgcg	actccttccc	cgactgtatc	gacggctccg	58620
acgagctcat	gtgtgggtgag	ccagcttctg	gcacggggaa	ggggcggtccg	ggctgggttc	58680
ccccaggaac	gtggagttaa	ggggaggaga	cgtgcctttc	cagcggggct	gggggctgtg	58740
tgggagactc	aggcggctgg	gaggtcctt	gcgggaggca	gggaagcctt	tcccagggca	58800

gcggccagga	ggacagactg	tgagctgtgg	gctcggcggc	tacagagtct	gcctcagtgg	58860
gcggggctga	tgggtgtccag	gtgcctgcag	cacgcaccca	cccacgggac	cttgcctgagc	58920
agcgtctgtc	aggcagcaag	attacccgag	ggctgcagtg	gtcctgttcc	ctggcagctt	58980
actgtctggc	tgaggaggag	tgatgttcac	atatgcacac	atgtcatgtg	cacacacatg	59040
tacatgacaa	catcccacat	gctcctcaaa	tagcatgacc	tgtacagtca	cggatatagg	59100
gcctagggga	taggaggcca	agacagtcag	ggaagacttt	ccagaggcag	tggctcctga	59160
aaggctgtct	gattcaggca	ggaaggaggc	tgagttcaga	taggaagtag	caatgagtca	59220
ttgtgtctgg	ggacatggcc	actccttcgc	tgcagaggga	cctgggctga	gagctcctct	59280
cttatggctg	cagtcgggag	agaagtctgt	tggggggaga	agggggcttc	ctcaagggac	59340
tccctgtgcc	ctttggcacc	ttcgtgccag	gtcaggcttg	aggcctgaag	gcagtgggtg	59400
gggccaccaa	gggtgcgctc	ctctgtctgg	caagttccca	gtctgacggg	cctgtgccgt	59460
gggccccagc	tgtggggggc	ctgttgatgc	gcagccaggc	ctcgccgcca	gagcccgcac	59520
gcttccattc	cgtgacttc	atcgacgcc	tcaggatcgc	tgggcccggc	ctgtgggaga	59580
gtgaatgtgg	cttttgccaa	agttgagtct	ggagcctgga	aacttcccta	tgggcagcct	59640
tgatagtggg	gtggcccaag	gagcccaccc	agccgaccct	gccccctccg	tggctgggtg	59700
gcggcaccag	gggtgcctg	gctttgtctg	ttcaccaaca	tcaccggggc	tggccagggc	59760
gcgctcactt	ctgccaccac	cgagggccct	gggcgaaggga	gtgaatacca	ggctgccttg	59820
gcagggatgt	gttgagggtc	gtggggagtc	ggacagcggc	gggggtcaga	ggaggaggag	59880
ggtgcaccgt	gcaggctgaa	gggccacggt	accctgaggt	tggccaggct	ccccaggcct	59940
agcctcccag	ctccccact	ttctccccac	cctccaccag	tggcaaagcc	agccccctca	60000
gggcgcacgg	tgtctgcccc	caaggagggc	ccattccggt	gggggttaatg	ttggccacct	60060
ctttctgttt	gtctctggca	gaaatcacca	agccgcctc	agacgacagc	ccggccccaca	60120
gcagtgccat	cgggcccgtc	attggcatca	tctctctct	cttcgtcatg	gggtggtgtct	60180
attttgtgtg	ccagcgcgtg	gtgtgccagc	gctatgcggg	ggccaacggg	cccttcccgc	60240
acgagtatgt	cagcgggacc	ccgcacgtgc	ccctcaattt	catagccccg	ggcgggttccc	60300
agcatggccc	cttcacagggt	aaggagcctg	agatatggaa	tgatctggag	gaggcaggag	60360
agtagtctgg	gcagcttttg	ggagtggagc	agggatgtgc	taccccaggc	cctcttgcac	60420
atgtggcaga	cattgctaata	cgatcacagc	attcagcctt	tcccactgag	cctgtgcttg	60480
gcatcagaat	ccttcaacac	agaggcctgc	atggctgtag	caaccacccc	tttggcactg	60540
taggtgtgga	gaaagctcct	tggacttgac	cttcatattc	tagtaggaca	tgtgctgtgt	60600
tgtccacaaa	tcctcatgta	ccctagaaat	gaatgtgggg	gcggctgggc	tctctccaga	60660
gctgaaggaa	tactctgtga	ccatacagca	gctttgtctt	gagtgcagct	gggatttgtg	60720
gctgagcagt	tacaattcct	acgtggccca	ggcaccaggga	acgcaggctg	tgtttgtaga	60780
tggctgggca	gccgcaccgc	agagctgcac	catgctgggt	tgtatcacat	gggtgaccat	60840
ggtatgtcta	agaagggtgga	gtccctgtga	ggtctgcagg	tgccccaca	gctccaggcc	60900
accttaggga	ttgcctctgc	ctgcccagcc	ctgagttccc	tctccccgt	cctgtcccac	60960
tgtcacccca	agccggccctc	attgggagcc	tgttggatgg	cagggtatag	atgtaacctg	61020
attctctctg	gggagtcggg	ttatctggct	tcacagagc	tcctaggagc	ccacagtgggt	61080
ggcaccatca	cagtgcgagc	agcccccaga	gaacgcggcc	ctgtctgttc	ctggcggtgct	61140
ctgtgctgcc	ccgcctgggt	tccctgcccc	agtcgcaggc	cccttggagg	aggtagccatg	61200
tgtctcccgt	ttcacagatg	agccccgggg	agctcactct	agtagtggcc	agagaggcct	61260
gcggctcagg	gagcggggca	catttccaac	aggacacacc	gcccctggct	gagtctcggtg	61320
ggtagtggga	gcagaggaga	gcgccttatg	tctgtggggc	ggcttggctg	agcctggaag	61380
ccacctgacc	tcccccgctc	cttccctgcc	aggcatcgca	tgcggaaagt	ccatgatgag	61440
ctccgtgagc	ctgatggggg	gccggggcgg	ggtgccccct	tacgaccgga	accacgtcac	61500
aggggcctcg	tccagcagct	cgtccagcac	gaaggccacg	ctgtacccgc	cggtgagggg	61560
cggggccggg	gaggggcggg	gcgggatggg	gctgtggggc	cctcccaccg	tcagtgtctg	61620
ccaccggagg	cttcccgggt	tcctgggggc	tgtgccaccg	cctctgaggc	atgcttgctt	61680
tcttcccctt	tcaaacctct	ctgcttcctt	ctttaatgac	attgttgatt	gtggataatc	61740
tgaaaactac	acaaaaatat	aaagagccaa	aatctcacc	aaatccacct	cctagagtgg	61800
ctgttgggct	ccgtcagcat	ccaggcggcc	gtctgtgttc	cgcacggccc	agcccatcga	61860
tagccgcctg	caccaggcct	gtctgcctc	tgtgagcctc	cccacagggt	tccctccaca	61920
aacacccctg	tctcccaccc	agggctggct	gcttccctga	aaacagctgg	atggttttgt	61980
gcatgacaga	caaacacagg	gtgattttcg	tggctaaaaat	actccctgga	gcttttgcca	62040

gggtgagggg	ctggctccag	ctgagccacg	ccttgagtga	aatgactgtg	aggagaataa	62100
actgccgctg	ccctccagga	tcactggggc	tggctgggga	gaacccccgt	ttctgggagc	62160
acagtcccag	gatgccaagg	cgagcttggt	gccgagatgt	gaactcctga	gtgtaaacag	62220
cgggggctga	cttgacatgc	tttgtatgct	tttcatttgt	tcctgcagct	gtatgccctt	62280
aaggtgagtc	cagccccctt	ctgcttctct	tggggcctcg	ccagtgagcc	ccaccttget	62340
ggggctgggt	cctcctgccc	ttctgggtat	ccctcacatc	tggggctctg	tcttcttggt	62400
ttatttttct	tttttttttg	agaoggagtt	tcacttttgt	tgcccaggct	tcagtgcaat	62460
ggtgtgatct	ctaggctcac	cgcaacctct	gcctcccagg	ttcaagcagt	tctcctgcct	62520
cagcctccct	agtagctggg	attacaggca	tgtgccacca	cgcccagcta	atthttgtatt	62580
tttagtagag	atggggtttc	tccatgttgg	tcaggctgat	cttgaaactcc	ctacctcagg	62640
tgatccgccc	accttggcct	cccaaagtgc	tgggattaca	ggcgtgagcc	accgcacctg	62700
gcctttttct	tttcttttct	tttctttttt	ctgagacagg	gtctcgctct	gtcaccaggg	62760
ctggagtgca	atgggtgtcat	catggctaac	tgcagcctct	accttctagg	ctcaagcaat	62820
cctcccctct	cagccccctaa	gtagctagga	ctgcacgcat	gcacccccat	gcccagctaa	62880
tattttacatt	ttttgtagag	atgaagtttc	actatatgtc	ccaggctggg	ctccaactcc	62940
tggactcgag	cgatcctcct	gcctcggcct	ccccagggtgc	tgggattaca	ggcgtgagcc	63000
accgtgcctg	gcctggggta	ttgtcttctt	attggcacctg	actgtgggtg	gccctgggaa	63060
ggaagttagca	gaagagggtt	cttcttggtt	tcctggacag	taactgagtg	ttctggaggc	63120
cccagggcct	ggctttgttt	agggacaaag	ggaactggta	accagaagcc	gagagtttaa	63180
acaccactg	cccttcttcc	ctgctcctgc	tgctgcaacc	cagcttaacc	agccaggagt	63240
gctaggaacc	caagcagggc	ccccgagcac	acagcaggca	gctcacgaat	tctcttttcc	63300
tgttctccct	tgggagctgg	gaggatctta	atcaggcaat	aagagatggc	actgagcagc	63360
cagctaattt	tttaaatacac	tttattgttt	aaccatatga	ctcaccact	taaaaaagg	63420
tacagttcag	tgggttttag	tgtattcaca	gatgtgtgca	acctcacca	cagttaattt	63480
tagaacattt	tcctgcccc	aaaagaaact	ctgcatgaag	ccagctgttt	ttaaattagc	63540
aaagtatttt	tgcacctctt	aaatatatgt	tcattgttaca	aaattcaaaa	gatacagaag	63600
agtctgcagt	ccaaagagac	tccgccccca	tgacgccaag	caggactccc	tgggaggcat	63660
ggcctcctgc	agtgtgtttc	ttctatgtcc	ccccaggggt	catctgtaca	tatgcaagca	63720
tacaagagcg	tggactttgt	tttccaagcc	agaagataat	tgtagattta	tgtgcagttg	63780
tgagaaagag	cacagaccca	tttatcctct	gcctggtttc	ccccagtget	gcctgccatc	63840
ttgcatgact	tccattccta	tcataagcaa	gacactgata	acgattcttt	caccttatct	63900
agattgacat	aagtgttttt	tgtttgttct	tgagacaaac	ttcctctgtc	accagtgagg	63960
agtgcagtgg	cacaatcaca	gctcactgca	gcctcaaact	cctgggctca	agcgattctc	64020
ctgcctcagt	ccctcaagt	agctcagatg	gcagggtgtc	accatcatgc	caggctaatt	64080
tttaaaattt	ttgtggagg	gaggcctcac	taaatttctt	gggctagtct	tgaactcctg	64140
agctaaagtg	atcctcctgc	ctcagcctcc	caaagtggta	ggattacagg	catgagccac	64200
tgcgcctggg	ctgacatatg	tgttttcgta	agccgaaaag	atagcatctg	aagagtcaac	64260
attgagcctt	gccttttgc	gctaattgatg	tataaaagct	gctgttctga	gcatttcgga	64320
ggctcccagc	tgcctgtgtc	accctgccta	gagctctacc	gtaaccctac	tccgggagga	64380
ggtgctattg	ttttcctcat	tttgcaacaa	ggaggctgaa	gaactgagca	tgaaccactg	64440
gcctgggtcg	ttcggttggg	aggcagtg	gccaggccat	ccaactcaca	accaccttct	64500
actctgcttc	ccccgcaccc	tgaagtttgt	tctgttttga	ggacacagcc	gtcacattct	64560
tgggtggctga	acagcactcc	ttgtcagggtg	tggctggggc	cccactggag	ggcatcatgg	64620
tcctctctcc	tgtgcgggtt	gaaccttggc	tgtttcaacc	actcctgcca	agtggccctc	64680
tgaaggggac	agtccatctt	ttctcagcag	agggccacac	tggcaaaacg	gtccctggca	64740
ccctttctct	ccacctgtct	aatatagagt	aaaaatggta	tcattgttaag	atcttcatth	64800
atattttatt	tatcatgaat	gatgtaagca	tcattttgtg	tgtttaagaa	cctttggggc	64860
cagcgtgatg	gcttgcagct	gtaatctcag	cacttttagga	ggctgagatg	agcggatcac	64920
ttgaggccgg	gagtttgaga	ccagcctggc	caacatggag	aaacccccgt	tctagtaaaa	64980
atttaaaaaat	tagccgggta	tgggtgatccc	agctacttgg	gagctctgaag	catgagaatt	65040
gcttgaacat	gggaggcgga	gggttcagtg	agccgagatc	gcgccattgc	actccagcct	65100
gggcgacaga	gcgagactct	gtctcacaaa	aaaaaaaaaa	aaagaaaaga	aaagaaatta	65160
tcaatctcct	cttttatggc	atatatatat	atatatatat	atatatatat	ttatttccct	65220
ttcttggtta	tgttcataaa	ggcctccctt	gctctgatca	taaaaaacaa	cttattttca	65280

cactctctct	cttttttttt	tgagacagag	ttttgctcct	gttgcccagg	ctggagtgca	65340
gtggcgcaat	ctcagctcac	tgtaacctcc	gcctcccggg	ttggagtgat	tctcctgcct	65400
taccttcccc	agtagctggg	attataggca	tgcaccacca	tgcctggcta	atthttgtact	65460
tttagtagag	acgggggttt	ctccatgttg	gtcaggctgg	tctcgaactc	gcgacctcag	65520
gtgateccacc	cacctcgccc	tcccaaagtg	ctgggattac	agacgtgagc	caccatgccc	65580
agcccacact	ctctttctta	acgtcctcct	cctttcgttt	tacgttcaca	tctttaattc	65640
ttctgggatg	taattagatt	tgatgagcaa	ggtgggcatc	cagcttgttt	cttggctgat	65700
ggcttatggg	tggcgtgaat	tagtcggggg	ctatcaggag	gcagaaactc	tatgagaatt	65760
tgaacagaga	aagtcccgtc	tacaggctta	ttaccaggga	ctggaatagc	agaaattgaa	65820
cagtgagatg	tacagagaac	tctaagaatg	caggaatagg	ccaggcatgg	tggtcacac	65880
ctgtcatccc	agcacttttg	gagaccaagg	cgggtggatc	acctgaggtc	aggagttcga	65940
gaccagcctg	gccaacatag	tgaaccccca	tctctactaa	aaatacaaaa	aaattagctg	66000
ggtgtggtgg	cgcatgcctg	taatccagc	ttctcgggag	tctgaggctg	gagaatcact	66060
tgaacctggg	aggcagaggt	tgtagtgagc	cgagatcatg	ccattgtact	ccagcctggg	66120
caacaagagc	gagactcagt	caaaacaaca	acaacgcagg	aatagcagat	gagccgaggt	66180
ggggcctccc	cagccccccac	ccccaccccc	gcaccctggg	ccgagatcca	gtcctctttg	66240
aatagggcct	gggcgtgggt	cacgggacat	ctgagacatt	gccgaggcgc	tgcactgggt	66300
gatcttgcca	gaagtctgcc	cagtgcagat	ttgggcagaa	tctcaaactg	ccttgggatg	66360
taggagagaa	accaggcctg	gtcaagttca	tgggaagagg	tggaaacaga	ccccataggc	66420
tggggccttg	gcagctgtag	gaagccctct	ctgctgcctc	cctgcctgct	ctctgctttg	66480
aagcatcttc	cccagtgcct	ccagtctcat	gccctctcaa	cgttgggggc	aaatcctgag	66540
gaatacccag	actggctctc	tgggccaaag	aggaccctct	ccagaaagag	cagggcccag	66600
tgcggcttcc	taaagggcag	gggaagggcc	tggccactcc	ccagaggcta	ctcaccagcc	66660
atcaggatag	cccaggaag	caggccttct	cgagcccatt	ttattacttt	atthttattat	66720
tttattttaat	tttaaattta	tttttttgaga	cagagtctca	ctctgttgcc	caggctggag	66780
tgcagtgggtg	cgatctcaac	ccactgcagc	ctctgcctcc	agggttcaag	ggattctccc	66840
acctcagcct	cccaagtagc	tgggattaca	ggtgcccgc	accacaccgc	gctaattttc	66900
atatttttag	tagagacgag	gtttcaccat	gttggccagg	ctggtctcga	actcctgacc	66960
tcaagtgate	cgcccgcctc	ggcctcccaa	agtgcctagg	caagcccatt	ttaaagttga	67020
agaaactgag	gctgaggtaa	attccctccc	cagggatcct	gctgcagcca	gaaggtggta	67080
aaacaggact	tcacccgggt	ctgtctggcg	tgaagggcag	tgttcttgta	ccaccctagg	67140
gggcctgaga	gaactgagtc	cctcgggcag	aactgacagt	tctgttccca	ttattccgca	67200
ggggctcgga	tctggctgta	tgttttcag	gatggccttg	gagaccacac	taagccctac	67260
accctttggg	aagctgcagt	ttgggttggg	gtgcctcag	tggcacttgt	ggaaggtgca	67320
gacctgtgtg	ggtgtgtggg	cccagggccc	ctggctccct	cctccctttg	tagggctggg	67380
tgtgtgctgc	ctggacctgg	ggggcacgtt	cacgtggtga	atthtgtctat	ttactatccc	67440
cgctttgggg	ctgggtgccag	cacaggccct	tgtgaagggg	gtgcctttgt	ctggagtggg	67500
actgtggccc	ctccctcagc	gtggtgactt	ctgtgtcagg	gcttcagcag	ggacgcagag	67560
cccctgagtg	ttcggaacaa	gggcgtcatt	gcaggagtta	gactgtgtgt	gatggaggga	67620
ggagggggcag	gaggaaaggt	cagaaggaga	gttccctggg	aggtccctga	ggagcctggg	67680
gaggtgctaa	ctggtgtgga	ggacactcag	ggcctgtggg	gacatctcct	actgctgggg	67740
gccagccaca	aagggaactg	gccgaagtcc	tgtcccgcgc	ttcacagccc	agcatctggg	67800
cacaaggcag	gtacttggaa	gggcgcgggc	acctggggca	aaagtgcctg	ggttcccttt	67860
gcctttcact	gagatgacct	tccggggcag	tggctgtctg	ctccctcctc	gtccccaggg	67920
tttgccaaact	ggccagagga	aggggtcctg	ggaagcaggg	gggcccagaag	ccctctctgc	67980
aaggaaagcc	cgaggggtgt	gggaggaagg	aaggaaatgcc	caggctggcg	aggtctctaa	68040
tcacccctggc	ttggctctcc	tcagatcctg	aaccgcgcgc	cctccccggc	cacggacccc	68100
tcctgtaca	acatggacat	gttctactct	tcaaacattc	cggccactgc	gagaccgtac	68160
aggtaggaca	tcccctgcag	ccctccatgg	ccattgggtt	cccggccagcc	cgtggtggag	68220
gggcctaata	cccatgccac	tgatgagggg	aggtattctg	ggtgctagtg	ggcaggtgcc	68280
ggggccagcc	ctgcctccct	ctgctctgcc	aaccacacta	ggctgcctcc	ccagacaagc	68340
tcagcgggca	ctgcatgttg	ggttcagaaa	tcagcagaac	tccacgttct	gagctgctct	68400
tcaagttgct	cctatggggg	ttacttttaa	gctgggaaat	ggctgtggcg	tccagggggc	68460
gggggcttgg	gctccaaact	ctgactgtgt	gtttgagtc	ggctgtggaa	acctagccat	68520

tgagatgccc	cctcttggtg	gctctgtect	cttaggatgg	gacaagtctg	tgaaggctgc	68580
tgcagcacc	accgtagacc	cctaatacgtg	tgacgtcacc	aggatggtcc	gggctgctca	68640
cttgccacag	tggcctgttt	gagcccgga	agccaacggg	gctgctcagc	tggacaccag	68700
ccccccgagc	tgcccatgtt	ggggtcacag	gccccacctc	cctgggttggg	gaggggcaac	68760
tgagagtgtg	gagaggtggg	accaggtgt	gctgggtctcc	gcaggggctg	gatcagagcc	68820
tgggatgggc	agggtagacc	tcctgacctt	taacccagtg	gtgtcaggca	acgtggccca	68880
cccgccagcc	gcaccaggcc	ccacccccgc	aggtgaaggg	gtgggatagg	ctgggcctgg	68940
gccaggacac	ctctggacca	cgcatctctc	attgcttggg	tccttgagc	agcagggcct	69000
cccgagtgtg	gtgcgcctg	ccacctagt	gccatttcca	cgaactccca	ggcctggctg	69060
gggagccgga	actgcagcct	ccatttccac	cccactccgg	gtcgggccac	ctccctgatg	69120
cctcagtatt	atatcaaaact	gtcacagtct	gtcccacagc	cttacagacc	actgtctcca	69180
gaatggtcac	atocacactg	ggcagcccag	tctcgctagt	tcctcgctcc	acctcctgcc	69240
tttgctcatg	cccgctctgc	tctgggcccc	ccgcggaac	atcttcccc	cgcccgccgt	69300
ctgacctcac	agcagctggg	ccccaaagagg	agtatcctgt	cctgctgcac	ttttctcaac	69360
acccggtgtt	ggctgcacct	tcaccccat	tgcaggcccc	tctgtgacag	gacgggggct	69420
cctaaacaca	ccacagttcc	gagtctgaac	tcacacagtg	ggatgcggcg	tttctgggcc	69480
acagttgggt	gcaggtagcc	tctgggagga	tgggaggtca	ggagccatct	tgcgagttag	69540
gttgcttgaa	ctcaggatgg	aagtgttccg	ggcccatagg	ttgctgtatt	agcctgttct	69600
cacgctgcta	ataaagacat	acccaagact	gggttaattgt	aaaggaaaga	ggtttaacgg	69660
actcacagtt	ccacctgcct	ggggtggcct	cagaatcatg	gtagaagaca	aggaggagca	69720
agtcacatct	tacatggcct	caggggaacag	acagcatgag	aaccaagcga	aaggggtttc	69780
cccttgtaaa	accatcaagt	ctagttagat	ttattcacta	ccacgagaac	agtatggggg	69840
gaaccacccc	catgattcaa	tcatctccca	ctgggtccct	cccacagcac	gtgggaatta	69900
tgggagtaca	attcaagatg	agatttgggt	ggggacacag	ccaaacccta	tcggttgcca	69960
acattttacag	taacagtgtt	aggtgaacag	ttgtccagtc	tcctgttttg	tcggacactg	70020
tttctagcac	cttcaggca	gaatctcatg	tatccttcac	tttcgaaatg	ggtactatth	70080
catccccact	tttatcaatg	agaaactaaa	gctcgaagag	gtcaagtaag	ttcctggcca	70140
aggtcagcta	gcaggctcta	gaggcctcgt	tctccttaga	ggcagccttg	ccagggccca	70200
ggcttggcag	gctgcagggc	aggtgcgggc	atgcccatgg	tagaggtggg	accattgagg	70260
ctcagagagg	gtaagtgatg	agccctggcg	acacagcggg	gtgggtccag	agtccggcct	70320
gcatcttctg	gagctggcca	gtggacaggc	ctttcccggt	cacagccccg	gggctgctgt	70380
gccaccagg	gcggtgtgtc	ctaccgaatc	ccactcctct	gtgtgtgtcc	ctttcaggcc	70440
ctacatcatt	cgaggaatgg	cgccccgcac	gacgccttgc	agcaccgacg	tgtgtgacag	70500
cgactacagc	gccagccgct	ggaaggccag	caagtactac	ctggatttga	actcggactc	70560
agacccctat	ccacccccac	ccacgcccc	cagccagtag	ctgtcggcgg	aggacagctg	70620
cccgccctcg	cccgccaccg	agaggagcta	cttcacatct	ttcccgcctc	ctccgtcccc	70680
ctgcacggac	tcactcctgac	ctcggccggg	ccactctggc	ttctctgtgc	ccctgtaaat	70740
agtttttaaat	atgaacaaag	aaaaaataat	atttttatgat	ttaaaaaata	aatataattg	70800
ggatttttaaa	aacatgagaa	atgtgaactg	tgatgggggtg	ggcagggctg	ggagaacttt	70860
gtacagtggg	gaaatattta	taaacttaat	tttgtaaaac	agaactgcca	ttcttttgtg	70920
ccctgtgtgc	atttgagttg	tgtgtccccg	tggagggaat	gccgaccccc	ggaccaccat	70980
gagagtccct	ctgcaccggg	gcgtccctct	gtccggctcc	tgcagggaag	ggctggggcc	71040
ttgggcagag	gtggatatct	cccctgggat	gcacccctga	gctgcaggcc	gggcccggctt	71100
tatgtgcgtg	tggcctgtgc	cgtcagaaag	ggccctgggc	ttcatcacgc	tgttgctgtt	71160
cgtcttccct	agattcttag	tccttttttt	tttttttttt	ttttgagacg	gagtctttct	71220
ctgtcatcca	ggctggagt	cagtgggtaca	atctcagctc	actgcaagct	ccgactccca	71280
ggttcaagt	agtctcctgc	ctcagcctcc	cgagttagctg	ggactacagg	tgcgcgccac	71340
cacacccgcc	cagctaattt	ttgtattttt	agtagagatg	gggtttcacc	atgttgccca	71400
ggatgatctc	gatctcttga	cctcgtgatc	cgccacacct	ggcctcccaa	agtgtggga	71460
ttataggcat	gagccactgt	accagctga	ctcttagtca	cttttaagaa	ggggactgtg	71520
ccttcatttt	tcactgggccc	ctgcagaata	tatgcctggg	ctctgggctc	ttctgaacct	71580
gtgttggtt	ccatctgacc	tctctgtgcc	agcccaaggc	tgtgtctctt	cctgagggca	71640
aggagcccca	tgactgcgtg	ttgactcgt	ggatggggct	gctgagccca	ctctgccaca	71700
ccacgtgccc	ctggcaggga	gggaatccct	gggtcctcac	aggaacagtc	agcaagccac	71760

acctgacgcc	tgctgtgggc	ccatccctgc	ggtgctggag	aagacagaca	aggcctggtc	71820
actgcctctg	caggggtccc	agtccgtgga	aggagacagt	aatctaggca	ttttcgggtg	71880
ggaagctgag	ctgttctcgt	gtcctgaagg	ccaggcggga	acagccgtct	tcagagggaa	71940
gggagaaaat	gcacatcgca	tcagtggaga	agggcctgac	ttccctcagc	atggtggagg	72000
gaggtcagaa	aacagtcaag	cttgagtatt	ctatagtgtc	acctaaata		72049

<210> 10

<211> 8705

<212> DNA

<213> Homo sapiens

<400> 10

ggactcaggg	gcagcagggg	ggtacaccca	tggttagtgg	gcggaccata	gggggtaatg	60
agagggtgaa	tcgatggaac	ctggggggaca	caatcgaagt	ggttccagag	tcgggctgta	120
ctaattaaag	agacggggca	gtggacaggc	attttcagtt	gactgcccag	ggagtgttct	180
gcccacagg	gaggatatgc	gtacagaatc	atactcgatc	agcatgagtc	caattcagac	240
cgtacatcag	tggagatatg	ggtccccga	tgactccgtg	gaacactgat	gtttgtgaca	300
ggggagtaca	gcaccagcca	tcagcaggcc	agtaaactcat	accggcctgc	gaaattggac	360
tcagaccogg	atccaccctg	accgacgtcc	caagccccc	ccccccacc	cccaccatgg	420
gccgagatcc	agtccctctt	gaataggggc	tggccgtggg	tcacggggaca	tctgagacat	480
tgccgaggcg	ctgcattggg	ggatccttgc	agaagtgtgc	ccagtgcaga	tttgggcaga	540
atctcaaaact	gccttgggat	gtaggagaga	aaccaggcct	ggtcaagtcc	atgggaagag	600
gtggaaacag	accccatagg	ctggggcctg	ggcagctgta	ggaagccctc	tctgctgcct	660
ccctgcctgc	tctctgcttt	gaagcatctt	ccccagtgcc	cccagttcca	tgcctctcca	720
acgttggggg	caaactctga	ggaataccca	gactggctct	ctggggccaaa	gaggaccctc	780
tccagaaaga	gcagggccca	gtgcggtctc	ctaaagggca	ggggaagggc	ctggccactc	840
cccagaggct	actaccagc	catcaggata	gccccaggaa	gcaggccttc	tcgagcccat	900
tttattactt	tattttatta	ttttatttaa	ttttaaattt	attttttgag	acagagtctc	960
actctgttgc	ccaggctgga	gtgcagtggg	gcgatctcaa	cccactgcag	cctctgcctc	1020
cagggttcaa	gggattctcc	cacctcagcc	tccaagtag	ctgggattac	agggtgccgc	1080
caccacaccc	ggctaatttt	catattttta	gtagagatga	ggtttcacca	tgttggccag	1140
gctggtctcg	aactcctgac	ctcaagtgat	ccgcccgcct	cggcctccca	aagtgttagg	1200
tcaagcccat	tttaaagtgt	aagaaactga	ggctgaggta	aattccctcc	ccagggatcc	1260
tgctgcagcc	agaagggtgg	aaaacaggac	ttcaccgggg	tctgtctggc	gtgaaaggca	1320
gtgttcttgt	accaccctag	ggggcctgag	agaactgagt	ccctcgggca	taactgacag	1380
ttctgttccc	attattccgc	aggggctcgg	atctggctgt	atgctttcca	ggatggcctt	1440
ggagaccac	ataagcccta	caccctttgg	gaagctgcat	gttgggttgg	ggtgccgtca	1500
gtggcacttg	tggaagggtg	agacctgtgt	gggtgtgtgg	gccaggggcc	cctggctcct	1560
tctcctcttt	gtagggctgg	ttgtgtgctg	cctggacctg	gggggcacgt	tcacgtgggtg	1620
aatttgtcta	tttactatcc	ccgctttggg	gctggtgccca	gcacaggccc	ttgtgaaggg	1680
ggtgcctttg	tctggagtgg	gactgtggcc	cctccctcag	cgtggtgact	tctgtgtcag	1740
ggcttcagca	gggacgcaga	gcccctgagt	gttcggaaca	agggcgtcat	tgcaggagtt	1800
agactgtgtg	tgatggaggg	aggaggggca	ggaggaaagg	tcagaaggag	agttcctggg	1860
aagggtccctg	aggagcctgg	tgagggtgcta	actggtgtgg	aggacactca	gggcctgtgg	1920
ggacatctcc	tactgctggg	ggccagccac	aaagggaact	ggccgaagtc	ctgtccccgc	1980
cttcacagcc	cagcatctgg	tcacaaggca	ggtacttgga	agggcgcggg	cacctggggc	2040
aaaagtgcct	gggttccctt	tgcctttcac	tgagatgacc	ttcggggcag	gtggctgctg	2100
cctccctctc	tgtccccagg	ttttgccaac	tggccagagg	aaggggtcct	gggaagcagg	2160
ggggccagaa	gccctctctg	caaggaaagc	ccgaggggtg	tgggaggaag	gaaggaaatgc	2220
ccaggctggc	gaggctctaa	gtcaccctgg	cttggctctc	ctcagatcct	gaacccgcgg	2280
ccctccccgg	ccacggaccc	ctccctgtac	aacatggaca	tgttctactc	ttcaaacatt	2340
ccggccactg	cgagaccgta	caggtaggac	atcccttgca	gccctccatg	gccattgggt	2400
tcccgcagc	ccgtgggtgga	ggggccta	ccccatgcc	ctgatgaggg	gaggtattct	2460
gggtgcta	gggcagggtg	cggggccagc	cctgcctccc	tctgctctgc	caaccacact	2520

aggctgcctc	cccagacaag	ctcagcgggc	actgcatggt	gggttcagaa	atcagcagaa	2580
ctccacgttc	tgagctgctc	ttcaagttgc	tcctatgggg	gttactttta	agctgggaaa	2640
tggtgtggc	gtcgaggggc	cgggggcttg	ggctccagag	tctgactgtg	tgtttgagtc	2700
cggctgtgga	aacctagcca	ttgagatgcc	ccctcttggt	ggctctgtcc	tcttaggatg	2760
ggacaagtct	gtgaaggctg	ctgcagcacc	caccgtagac	ccctaactgt	gtgacgtcac	2820
caggatggtc	cgggctgctc	acttgccaca	gtggcctggt	tgagcccggg	aagccaacgg	2880
ggctgctcag	ctggacacca	gccccccgag	ctgcccattgt	tggggtcaca	ggccccacct	2940
ccctggttgg	ggaggggcaa	ctgagagtg	ggagagggtg	gacccagggtg	tgctgggtctc	3000
cgcaggggct	ggatcagagc	ctgggatggg	cagggtgagc	ctcctgacct	ttaacccagt	3060
ggtgtcaggc	aacgtggccc	acccgccagc	cgcaccaggc	cccacccccg	cagggtgaagg	3120
ggtgggatag	gctgggcctg	ggccaggaca	cctctggacc	acgcattcct	cattgcttgg	3180
gtccctggag	cagcagggcc	tcccagatgt	ggtgccgcct	gccacctagt	ggccatttcc	3240
acgaactccc	aggcctggct	ggggagccgg	aactgcagcc	tccattttcca	ccccactccg	3300
ggtcgggcca	cctccctgat	gcctcagtat	tatatcaaac	tgtcacagtc	tgccccacag	3360
ccttacagac	cactgtctcc	agaatggtca	catccacact	gggcagccca	gtctcgctag	3420
ttcctcgtcc	cacctcctgc	ctttgctcat	gcccgtcctg	ctctggggcc	accgcggaca	3480
catcttcccc	cgcgccgcg	tctgacctca	cagcagctgg	gccccaaagag	gagtatcctg	3540
tctgtctgca	cttttctcaa	cacccgggtg	tggctgcacc	ttcccaccca	ttgcaggccc	3600
ctctgtgaca	ggacgggggc	tctaaacac	accacagttc	cgagtctgaa	ctcacacagt	3660
gggatgcggc	gtttctgggc	cacagtggg	tgcaggtagc	ctctgggagg	atgggaggtc	3720
aggagccaac	ttgcgagtca	ggttgcttga	actcaggatg	gaagtgttcc	gggcccattg	3780
gttgctgtat	tagcctgttc	tcacgctgct	aataaagaca	tacccaagac	tgggtaattg	3840
taaaggaaaag	aggtttaacg	gactcacagt	tccacctgcc	tggggtggcc	tcacaatcat	3900
ggtagaagac	aaggaggagc	aagtcacatc	ttacatggct	tcagggaaca	gacagcatga	3960
gaaccaagcg	aaaggggttt	ccccttgtaa	aaccatcaag	tctagtgaga	tttattcact	4020
accacgagaa	cagtatgggg	ggaaccaccc	ccatgattca	atcatctccc	actgggtccc	4080
tcccacagca	cgtgggaatt	atgggagtac	aattcaagat	gagatttggg	tggggacaca	4140
gccaaaccct	atcggttgcc	aacatttaca	gtaacagtgt	taggtgaaca	gttgtccagt	4200
ctcctgtttt	gtcggacact	gtttctagca	ccttccaggc	agaatctcat	gtatccttca	4260
ctttcgaaat	gggtactatt	tcatccccac	ttttatcaat	gagaaactaa	agctcgaaga	4320
ggtcaagtaa	gttctctggc	aaggctcagc	agcaggctct	agaggcctcg	ttctccttag	4380
aggcagcctt	gccagggccc	aggcttggca	ggctgcaggg	cagggtgcggg	catgcccattg	4440
gtagagggtg	gaccattgag	gctcagagag	ggtaagtgat	gagccctggc	gacacagcgg	4500
ggtgggtcca	gagtcggg	tgcatcttct	ggagctggcc	agtggacagg	cctttcccgt	4560
tcacagcccc	ggggtctgtg	tgcccaccag	ggcggtgtg	cctaccgaat	cccactcctc	4620
tgtgtgtgtc	cctttcaggc	cctacatcat	tcgaggaatg	gcgccccga	cgacgccctg	4680
cagcaccgac	gtgtgtgaca	gcgactacag	cgccagccgc	tggaaggcca	gcaagtacta	4740
cctggatttg	aactcggact	cagaccccta	tccaccccca	cccacgcccc	acagccagta	4800
cctgtcggcg	gaggacagct	gcccgcctc	gcccgcacc	gagaggagct	acttccatct	4860
cttcccggcc	cctccgtccc	cctgcacgga	ctcatcttga	cctcggcggg	gccactctgg	4920
cttctctgtg	ccctgtataa	tagtttttaa	tatgaacaaa	gaaaaaata	tattttatga	4980
tttaaaaaat	aaatataatt	gggattttta	aaacatgaga	aatgtgaact	gtgatggggg	5040
gggcagggct	gggagaactt	tgtacagtgg	agaaatattt	ataaaactta	ttttgtaaaa	5100
cagaactgcc	attctttcgt	gccctgtgtg	catttgagtt	gtgtgtcccc	gtggagggaa	5160
tgccgacccc	cggaccacca	tgagagtcct	cctgcacccg	ggcgtccctc	tgtccggctc	5220
ctgcagggaa	gggctggggc	cttgggcaga	ggtggatatc	tcccctggga	tgcattccctg	5280
agctgcaggc	cgggcgggct	ttatgtgcgt	gtggcctgtg	ccgtcagaaa	gggcccctggg	5340
cttcatcacg	ctgttgctgt	tcgtcttctc	cagattctta	gtcttttttt	tttttttttt	5400
ttttttgaga	cggagtcttt	ctctgtcatc	caggctggag	tgagtggtta	caatctcagc	5460
tactgcaag	ctccgactcc	caggttcaag	tgagtctcct	gcctcagcct	cccagtagtc	5520
tgggactaca	ggtgcgcgcc	accacaccgc	cccagctaatt	ttttgtattt	ttagtagaga	5580
tggggtttca	ccatgttggc	caggatgatc	tcgatctctt	gacctcgtga	tccgcccacc	5640
tcggcctccc	aaagtgtctg	gattataggc	atgagccact	gtacccagct	gactcttagt	5700
cacttttaag	aaggggactg	tgcccttcatt	tttactggg	ccctgcagaa	tatatgcctg	5760

ggctctggggc	tcttctgaac	ctgtgttggc	ttccatctga	cctctctgtg	ccagcccaag	5820
gctgctgctc	ttcctgaggg	caaggagccc	catgactgcg	tggtgactcg	ctggatgggg	5880
ctgctgagcc	cactctgcca	caccacgtgc	ccctggcagg	gaggaatcc	ctgggtcctc	5940
acaggaacag	tcagcaagcc	acacctgacg	cctgctgtgg	gcccacccct	gcggtgctgg	6000
agaagacaga	caaggcctgg	tcaactgectc	tgcagggtcc	ccagtccgtg	gaaggagaca	6060
gtaatctagg	cattttcggg	ggggaagctg	agctgttctc	gtgtcctgaa	ggccaggcgg	6120
gaacagccgt	cttcagaggg	aaggagagaa	atgcacatcg	catcagtggg	gaagggcctg	6180
acttccctca	gcatggtgga	gggaggtcag	aaaacagtca	agcttggtgc	tgggtgacag	6240
tgcatttaat	aatcaaaata	taggctgggt	acggtggctc	atgcctgtaa	tcccagcact	6300
ttgggagggt	gaggcagggtg	gatcacttga	ggccaggagt	ttgagaccgg	cctggccaac	6360
atggcaaaac	ctcaactact	aaaatacaaa	aactagccgg	gcgtggtggt	gcacgcctgt	6420
aatcccagct	acttgggagg	ctgaggcagg	agaattgctt	gaacctggga	ggcggagggt	6480
gcagtgagcc	gagattgtgc	cactgcactc	cagcctgggc	aacagagcaa	gactctgtct	6540
caaaaaaaaa	aaaaaaaaaa	gcaatacaaa	atacaaatat	cactttcact	aaaagaaggg	6600
atggaagacc	caaaaacaaac	agaaaacaaac	aaaatggcag	gagtaagtcc	ccacttatca	6660
ataataacat	tgactgtaaa	taggctaagc	tctgcaatca	aaagagtggg	ccaggagcgg	6720
tggctcacgc	ctgtaattcc	aacgctttgg	gaggtgagg	cggatggatc	atttgatgtc	6780
acgagtttta	agaccagcct	ggccaacaag	gtgaaacccc	atctgtacta	aaaatacaaa	6840
aattagccag	gcggtagtgg	cacgcacctg	taatcccagc	tacttgtgag	gctgaggcag	6900
gagaatcact	ggaggctggg	aagcggagggt	tgctgtgagc	caagatggag	ccactgcact	6960
cccacctggg	cgacagagtg	agatcctgtc	ttaagaaaaa	aaagagtgga	tgaatggatc	7020
aaaaaacaag	acccaacctat	ctcttgcata	caagaaacac	actttacctt	taaaaacaca	7080
ctaggccagg	tgtggtgggt	cacacctgta	atcccagccc	tttgggaggg	ctgactggca	7140
gatcacctga	ggccaggagt	ttcagaccag	cttgaccgac	atggcaaaac	cccatctctc	7200
ctaaaaatat	aaaaaaacaa	aaaaaagaaa	aaggctggaa	gtagtgatgt	gtgcctgtag	7260
ccccagctac	ttgggagggt	gaggcaggag	aattgcttga	atccgggaag	tggaggttgc	7320
agtgagccag	gatggtgccca	ctgcactcca	gcctgggtga	cagagcgaga	ccctgtcata	7380
aaaaaaaaaa	gaaaagaaaa	gaaaaacgag	aaaaacaaac	acaaaattag	tagaagaaaa	7440
gaaataataa	agatcagaac	aggccagggt	catgggcaca	gtggctcaac	tcctacctgc	7500
tcaggagttt	gagaccagtc	tggccaacat	ggcaaaaccc	catctctcct	aaaaatatga	7560
aaaaaaaaaa	ataggctgga	tgtggtgatg	tgtgtgtgcc	tgtagcccca	gctacttggg	7620
aggctgaggt	gggagaatca	cttgagccca	ggaagtggag	gctgcagcga	gtcatgaatg	7680
caccctgcac	tctagctggg	taactggagt	gagattctgt	ctcaaaaaag	caaagaccag	7740
agcagaaata	aatgaaatgg	aaatgaaggga	aacaatgcaa	aatgatacaa	aaagtttttt	7800
cgaaaagata	aacaaaatca	acaaaacctt	agccagatta	agaaaaaaag	agagaagacc	7860
caaataaata	aaatccgaga	ttaaaaagga	gacattacca	ctgataccac	agaaattcaa	7920
aggatcatta	gaggcaacta	tgtgcaacta	tatgctaata	aactggaaaa	cctagaagaa	7980
ctgggtaaat	ttctagacac	atacaacctt	tcaagattga	accatgaaga	aatccaaaac	8040
ctgaacaggc	cgggcacggg	ggcttacgcc	tgtaatccca	gcactttgga	aggcctgaga	8100
tcaggagttc	gagaccagcc	tggccaacat	ggtgaaaccc	catctctact	gaaaaaatat	8160
aaaaaattag	cgggcgtggg	ggcgggtgcc	tctaattgtc	gccactcggg	aggctgaggc	8220
aggaaaatca	cttgaacctg	ggaggcatag	gttgacgcga	gccgaggttg	caccactgca	8280
ctccagcctt	ggcgacagag	ccagactcca	tctcaaaaaa	attaaaaata	caaaaacctg	8340
aacagaccaa	taacaagtaa	tgcgatgaaa	actgtaataa	aatgtttccc	aacaaagaaa	8400
gcccaggaac	aaatggcttc	actgctgaat	tttaccaaaac	attttttttt	ttttgagacg	8460
gagtctcgct	ctgtcgccca	ggctggagtg	cagtgggtga	acctcggttc	gctggtaact	8520
tatgcctctc	aggctgcaag	tgattttcct	gcttcaggcc	ccccgagtg	ctggaaatta	8580
gatggtactt	gtcaacaag	gcctggctaa	atttctatat	ttccttcaag	tagaagatgt	8640
gcttccaaca	aaggttgggt	tacggctggc	ttctgaaaat	cttggatttc	aaggctcccc	8700
aaaag						8705

<210> 11

<211> 66933

<212> DNA

<213> Homo sapiens

<400> 11

tataatcaag	cgcgttccgt	ccagtcggt	gggaagattt	tcgatatgct	tcgtgatctg	60
ctcaagaacg	ttgatcttaa	agggttcgag	cctgatgtac	gtatttttgc	taccaaatac	120
agcaatagta	atggctctca	gtccccgtgg	atggaggagc	aaattcggga	tgcttgggga	180
agcatgggtc	taaaaaatgt	tgtacgtgaa	acggatgaag	ttggtaaagg	tcagatccgg	240
atgagaactg	tttttgaaca	ggccattgat	caacgcctct	caactggtgc	ctggagaaat	300
gctctttcta	tttgggaacc	tgtctgcaat	gaaattttcg	atcgtctgat	taaaccacgc	360
tgggagatta	gataatgaag	cgtgcgcctg	ttattccaaa	acatacgctc	aatactcaac	420
cggttgaaga	tacttcgtta	tcgacaccag	ctgccccgat	ggtggattcg	ttaattgcgc	480
gcgtaggagt	aatggctcgc	ggtaatgcca	ttactttgcc	tgtatgtggg	cgggatgtga	540
agtttactct	tgaagtgtct	cgggggtgata	gtgttgagaa	gacctctcgg	gtatggtcag	600
gtaatgaacg	tgaccaggag	ctgcttactg	aggacgcact	ggatgatctc	atcccttctt	660
ttctactgac	tgggtcaacag	acaccggcgt	tcggtcgaag	agtatctggg	gtcatagaaa	720
ttgccgatgg	gagtcgccgt	cgtaaagctg	ctgcacttac	cgaagtgtat	tatcgtgttc	780
tggttggcga	gctggatgat	gagcagatgg	ctgcattatc	cagattgggt	aacgattatc	840
gccaacaag	tgcttatgaa	cgtggtcagc	gttatgcaag	ccgattgcag	aatgaatttg	900
ctggaaaata	ttctgcgctg	gctgatgcgg	aaaatatctt	acgtaagatt	attaccgcgt	960
gtatcaacac	cgccaaattg	cctaaatcag	ttgttgcctc	tttttctcac	cccggtgaac	1020
tatctgcccg	gtcagggtgat	gcacttcaaa	aagcctttac	agataaagag	gaattactta	1080
agcagcagcg	atctaaccctt	catgagcaga	aaaaagctgg	ggtgatattt	gaagctgaag	1140
aagttatcac	tcttttaact	tctgtgctta	aaacgtcatc	tgcatacaaga	actagttaa	1200
gctcacgaca	tcagtttgct	cctggagcga	cagtattgta	taagggcgat	aaaatgggtgc	1260
ttaacctgga	cagggtctcgt	gttccaactg	agtgtataga	gaaaattgag	gccattctta	1320
aggaacttga	aaagccagca	ccctgatgcg	accacgtttt	agtctacgtt	tatctgtctt	1380
tacttaatgt	cctttgttac	aggccagaaa	gcataactgg	cctgaatatt	ctctctgggc	1440
ccactgttcc	acttgtatcg	tcggctctgat	aatcagactg	ggaccacggg	cccactcgta	1500
tcgtcggctc	gattattagt	ctgggaccac	ggtcccactc	gtatcgtcgg	tctgattatt	1560
agtctgggac	cacgggtccca	ctcgtatcgt	cggctctgata	atcagactgg	gaccacgggc	1620
ccactcgtat	cgtcgggtctg	attattagtc	tgggaccatg	gtcccactcg	tatcgtcggg	1680
ctgattatta	gtctgggacc	acgggtcccac	tcgtatcgtc	ggtctgatta	ttagtctgga	1740
accacggtcc	cactcgtatc	gtcgggtctga	ttattagctc	gggaccacgg	tcccactcgt	1800
atcgtcggtc	tgattattag	tctgggacca	cgatcccact	cgtgttgctg	gtctgattat	1860
cggctctggga	ccacggtccc	acttgtattg	tcgatcagac	tatcagcgtg	agactacgat	1920
tccatcaatg	cctgtcaagg	gcaagtattg	acatgtcgtc	gtaacctgta	gaacggagta	1980
acctcgggtg	gcgggttgat	gcctgctgtg	gattgctgct	gtgtcctgct	tatccacaac	2040
atthtgcgca	cggttatgtg	gacaaaatac	ctggttaccc	aggccgtgcc	ggcacgttaa	2100
ccgggctgca	tccgatgcaa	gtgtgtcgtc	gtcgacgagc	tcgcgagctc	ggacatgagg	2160
ttgccccgta	ttcagtgtcg	ctgatttgta	ttgtctgaag	ttgtttttac	gttaagttga	2220
tgcagatcaa	ttaatagcat	acctgcgtca	taattgatta	tttgacgtgg	tttgatggcc	2280
tccacgcacg	ttgtgatatg	tagatgataa	tcattatcac	tttacgggtc	ctttccgggtg	2340
atccgacagg	ttacggggcg	gcgacctcgc	gggttttcgc	tatttatgaa	aattttccgg	2400
tttaaggcgt	ttccgttctt	cttcgtcata	acttaatgtt	tttatttaaa	ataccctctg	2460
aaaagaaagg	aaacgacagg	tgctgaaagc	gagctttttg	gcctctgtcg	tttcctttct	2520
ctgtttttgt	ccgtggaatg	aacaatggaa	gtccgagctc	atcgctaata	acttcgtata	2580
gcatacatta	tacgaagtta	tattcgatgc	ggccgcaagg	ggttcgcgtc	agcgggtgtt	2640
ggcgggtgtc	ggggctggct	taactatgcg	gcatcagagc	agattgtact	gagagtgcac	2700
catatgcggt	gtgaaatacc	gcacagatgc	gtaaggagaa	aataccgcat	caggcgccat	2760
tcgccattca	ggctgcgcaa	ctgttgggaa	gggcgatcgg	tgcgggcctc	ttcgctatta	2820
cgccagctgg	cgaagggggg	atgtgctgca	aggcgattaa	gttgggtaac	gccaggggtt	2880
tcccagtcac	gacgttgtaa	aacgacggcc	agtgaattgt	aatacgactc	actatagggc	2940
gaattcgagc	tcggtacccc	gggatcctct	agagtcgacc	tgcaggcatg	caagcttctc	3000
ttgtgccggt	tgtacgctgt	cagggtcacac	tggtaggtta	ggcagggcac	agatgccccag	3060

agcagagggga	acttttccttg	gggattcaac	acgtgcaagt	cttaggggct	ggcaaatacct	3120
gccctcagct	agagaggggg	cttttatattg	agaccagaat	cacctgagca	tcctcctgtc	3180
cccagctgtg	tccagcctgt	ctgcagggac	atcctgagag	gaccaggctc	tccctcctc	3240
cacctgccta	agtgccactc	tgaacctctgt	ccacctgtgc	cgtggagggg	cgtgacctca	3300
agctgctcag	ccagcagcag	gcttggccct	ggggggcagc	agagacccag	gtggctgtgg	3360
ggtgggtgct	tcgtggcgtg	gttctgaaac	ttcgttggaa	gtgtgtggac	agtgccttgc	3420
ctgttctctg	tgggacccta	tttagaaacg	aggtctgagt	tactgggggt	catcactgtg	3480
ttctgatggc	ccagctgtgt	ggaggccgcg	gtgcagcccc	atccaaggag	ccaggggcct	3540
gggtctagcc	gtgaccagaa	tgcattgcccc	ggaggtgttt	ctcatctcgc	acctgtgttg	3600
cctggtgtgt	caagtggctg	tgaacctctg	tgttagctct	tgggtgttct	gaaagtgcct	3660
ccgggtctca	ggcctcagaa	ccagggtttc	ccttcatctc	ggtggcctgg	gagcatctgg	3720
gcagttgagc	aaagagggcg	attcacttga	aggatgtgtc	tggccctgcc	taggagcccc	3780
ccggcacggg	gctggggcct	gaagctgccc	tcgggtggtg	gagaggaggg	agcgatgaag	3840
tggcgtcgag	ctgggcagga	aggggtgagcc	cctgcaagggt	gggcatgctg	gggacgctga	3900
gcagcatggc	cagcagctgg	gtctgcagcc	tggtagcccg	cgggacttgt	ggttggggct	3960
ggtttgtggc	caggagaggg	gctggcagga	gacaaggggg	actgtgaggg	agctcccacc	4020
cagcagctga	agcccaatgg	cctggctgtg	tggctctcag	ctgctgcat	aacctctcag	4080
tgttccagtt	ctctcatttg	taaaatgagg	aaacaaacag	tgccagcctc	ccagagggtgt	4140
catgaggatg	aacgagtgac	catgtagcat	gggctgggtg	cgtgtcacct	aacatcacca	4200
gcctttgcaa	ggagagccct	gggggccttg	ctgagtattt	cccttgcccg	gcccacccca	4260
ggcctagact	tgtgctgtgt	gcaggccctt	gacccttgac	cccattgcac	ctgtctccac	4320
aggagccgag	gaggtgctgc	tgttggcccg	gcgagcgac	ctacggagga	tctcgctgga	4380
cacgccggac	ttcacccgaca	tcgtgctgca	ggtggacggc	atccggcacg	ccattgccat	4440
cgactacgac	ccgctagagg	gctatgtcta	ctggacagat	gacgaggtgc	gggccatccg	4500
cagggcgtac	ctggacgggt	ctggggcgca	gacgctggtc	aacaccgaga	tcaacgaccc	4560
cgatggcatc	gcggctgact	gggtggcccc	aaacctctac	tggaccgaca	cgggcacgga	4620
ccgcacagag	gtgacgcgcc	tcaacggcac	ctcccgcaag	atcctggtgt	cggaggacct	4680
ggacgagccc	cgagccatcg	cactgcaccc	cgtgatgggg	taagacgggc	gggggctggg	4740
gcctggagcc	agggccaggc	caagcacagg	cgagaggagg	attgacctgg	acctgtcatt	4800
ctgggacact	gtcttgcac	agaacccgga	ggagggtctg	ttaaaacacc	ggcagctggg	4860
ccccaccccc	agagcggtga	ttcaggagct	ccaggggcgg	gctgaagact	tgggtttcta	4920
acaagcaccc	cagtggctcc	gtgctgctgc	tgggtccatg	cgtagaaagc	cctggagacc	4980
tggagggagc	cctttgttcc	cctggcttca	gtttcctcat	ctgtagaatg	gaacggtcca	5040
tctgggtgat	ttccaggatg	acagttagtg	cagtaagggc	agcctctgtg	acactgacca	5100
cagtacaggg	caggcctctt	tttttctttt	tttttttttg	agatggagtc	tcactctgtc	5160
gcccaggctg	gagtgcagtg	gtgtgatctc	agctcactac	aacctctgcc	tcctgggctc	5220
aagtgattct	cctgcctcag	cctcctgagt	agctgggatt	acagggtgct	gccactgtgc	5280
ttggctaattg	tttgtatttt	tggtagagat	gggttttcac	cgtcttggcc	aggctggctg	5340
caaactcctg	acctcagggt	atccacctgc	ctcagcctcc	caaagtgtct	ggattacagg	5400
catgagccac	cacgcccggg	caggccaggc	ctcttttgaa	cactttgcac	accatgggtc	5460
ttttcatcca	ggggggtagg	tacagttgta	cagttgagga	cactgaagcc	cagagaggct	5520
cagggaacttg	cccagggtca	cacagcagga	tgtggcagggt	gtggggctgg	gcctggcagc	5580
gtggctccag	ctttccagca	tagaaatctg	tgaagcaga	tagtttgtcg	gtcggtaggg	5640
gagactttct	gagacccgcc	ccagcggctc	agagggtagt	agccaggggc	cttcctgggg	5700
gctcataacc	cagaacactg	aatgggaaaa	ccctgatgga	ggaggcgcag	tggagctgtg	5760
ggtgccgatg	ggaagtccca	gaggagctgg	gaggtcagta	gcggtgctgc	cctctgtgga	5820
gcacttagtg	ggcaccagggt	gtgtttccag	gttcatggcc	ctgggacctg	aagctcagaa	5880
ggtgaagtaa	cttgcccagg	gcacccgtcg	ggcagcggcg	ggcagaggat	ttgtgggctg	5940
tggagcctgt	gctcgtggcc	cagccctggg	ggttgtgagt	gtgctggccg	gggagctttt	6000
cctgcaagtg	gactggtgtc	taggagccag	catgtcaggc	agcaggcagc	gggagtgcag	6060
caggcagcgg	gagcacagca	ggcagagggc	ggggctcgag	cagccatccg	tggaccctgg	6120
ggcacggagg	catgtgggag	agggctgctc	catggcagtg	gctgaagggc	tgggttgtgc	6180
cccaggaggg	gtggatgagg	gtaagaagtg	gggtccccag	gggctttagc	aagaggaggc	6240
ccaggaactg	gttgccagct	acagtgaagg	gaacacggcc	ctgaggtcag	gagcttggtc	6300

aagtcactgt	ctacatgggc	ctcgggtgtcc	tcactctgtga	aaaaggaagg	gatggggaag	6360
ctgactccaa	ggccccctcct	agccctgggtt	tcatgagctct	gaggatccca	gggacatggg	6420
cttggcagtc	tgacctgtga	ggtcgtgggg	tccagggagg	ggcaccgagc	tgggaagcggg	6480
aggcagaggg	gctggccggc	tgggtcagac	acagctgaag	cagaggctgt	gacttggggc	6540
ctcagaacct	tcacccctga	gctgccaccc	caggatctgg	gttccctcct	tggggggccc	6600
cagggaaaca	gtcacctgtc	ctttgcatag	gggagccctt	cagctatgtg	cagaagggtc	6660
tgtctgtccc	cttcctccct	ctaggtgtct	agctcctcca	gcccactagt	cagatgtgag	6720
gctgccccag	accctgggca	gggtcatttc	tgtccactga	cctttgggat	gggagatgag	6780
ctcttggccc	ctgagagtcc	aagggtgtgt	gtggtgaaac	ccgcacaggg	tgggaagtggg	6840
catccctgtc	ccaggggagc	ccccagggac	tctggtcact	gggcttgccg	ctggcatgct	6900
cagtcctcca	gcacttactg	acaccagcat	ctactgacac	caacatttac	aaacaccgac	6960
attgaccgac	accgacattt	accgacactg	acatttacca	acactgttta	ccaacactga	7020
catctactga	cactggcatc	taccaacact	gacatttacc	gacactgaca	tttaccaaca	7080
ctatttacca	acactgacat	ctactgacat	tggcatctac	caacaccaac	atttaccgac	7140
accaacattt	accaacactg	aaattttaccg	acaccgacat	ttaccgacac	cgtttaccaaa	7200
caccgacgtt	taccgacacc	gacattttacc	gacactgata	tttaccaaca	ctgacatcta	7260
ctgacgctgg	catctactga	caccgatgcc	agcatctacc	aacaccgaca	tttaccaaca	7320
ctgacattta	ctgacactga	tatctactga	cactggcatc	tactgacacc	aacattttacc	7380
aacaccagca	tctaccaaca	ccgacattta	ccaacaccag	cattttacca	caccgatgtt	7440
taccaacgcc	gacgtttacc	gacgccagca	tctaccaaca	ctgacattta	ccgacaccga	7500
catttaccga	cactgacatt	tactgacact	gacatttact	gatactggca	tctaccgaca	7560
ctgatattta	ccaacgccag	catctactga	cactgatgtt	taccaacacc	gacattttacg	7620
agcaccgaca	tttactgaca	ccaatattta	ctgacatcaa	cattttagcca	tgtgatgggg	7680
gccggcttgg	gggcaggcct	tgtctttggc	actgggggatg	ctgcagagac	cagacagact	7740
catgggggtca	tggactttctg	cttctttctcc	agcctcatgt	actggacaga	ctgggggagag	7800
aaccctaata	tgcagtggtc	caacttggat	gggcaggagc	ggcgtgtgct	ggtcaatgcc	7860
tccctcggtt	ggcccaacgg	cctggccctg	gacctgcagg	aggggaagct	ctactgggga	7920
gacgccaaaga	cagacaagat	cgaggtgagg	ctcctgtgga	catgtttgat	ccaggaggcc	7980
aggcccagcc	accccctgca	gccagatgta	cgtattggcg	aggcaccgat	gggtgcctgt	8040
gctctgctat	ttggccacat	ggaatgcttg	agaaaatagt	tacaatactt	tctgacaaaa	8100
acgccttgag	agggtagcgc	tatacaacgt	cctgtgggtta	cgtaagatgt	tatcattcgg	8160
ccaggtgcct	gtagacacag	ctacttggag	actgaggtgg	gaggatcgct	ggagtccaaag	8220
agtttgaggc	cagcccgggc	aaaggggaca	caggaatcct	ctgcactgct	tttgccactt	8280
actgtgagat	ttaaattatt	tcacaataca	aaattaagac	aaaaagttaa	tcacatatcc	8340
actgccttgc	ttaagacaga	aaacatgggt	gttgttgaag	ccagaggcag	ctgctggcct	8400
gagtttggtg	attggttcct	aagcagttga	aggcagtttt	gtttttccat	agatgtctgt	8460
tctccctttg	ctgggtgcag	cctcgccctg	ctgctgtggt	cgggtttcag	tggcctcgtc	8520
ccgtggacgc	agcctcgccc	tgccgctgtg	gtcgggtttc	agtggcctcg	tcccgtagac	8580
gcagcctcgc	cctgtcgtgt	tggtcgggtt	tcagtggcct	cgtcccgtgg	acgcagcctc	8640
gccctgccgc	tgtggtcggg	tttcagtggc	ctcgtcccgt	ggacgcagcc	tcgccctgcc	8700
gctgtggtcg	ggtttcagt	gcctcgtccc	atgggcgtgc	tttggcagct	ttttgctcac	8760
ctgtggagcc	tctcttgagc	ttttttgttt	gttgtttgtt	tttggttgat	tttggttgat	8820
tgtttgtttt	tgttgtcgtt	gttgttgccc	aggctggagt	gcagtggcgc	gatctcagct	8880
cactgaaacc	tctgcctcct	tgggttcattg	ccattctcct	gcctcagcct	cccacatagc	8940
tgggattaca	agtgcgcgcc	accacgcctg	gctaaatttt	gtatttttag	tagacagggg	9000
gtttcaccat	gttggtcagg	ctggctctgga	actcctggtc	tcacatgatc	cacctgcctc	9060
ggcctcccaa	agtgttggga	ttacaggcgt	gagccaccgc	gcccagcctc	tgttgagcat	9120
atthtgaggt	tctcttgggt	ccagtatat	gtacatgtgt	ccccatcgca	ccatcgtcac	9180
ccattgaggt	gacattgggt	cctctcctcg	gggtggatgc	ctccctctgt	ttccagcaac	9240
ttctgaagga	ttttcctgag	ctgcatcagt	ccttgttgac	gtcaccatcg	gggtcacctt	9300
tgtctcctc	agggctccca	ggggaggccc	gaatcaggca	gcttgcaggg	cagggcagga	9360
tggagaacac	gagtgtgtgt	ctgtgttgca	ggatttcaga	ccctgcttct	gagcgggagg	9420
agtctcagca	ccttcagggt	ggggaaccca	gggatggggg	aggctgagtg	gacgccttct	9480
ccacgaaaac	cctaggagct	gcaggtgtgg	ccatttcctg	ctggagctcc	ttgtaaatgt	9540

tttgtttttg	gcaaggccca	tgtttgcggg	ccgctgagga	tgatttgcct	tcacgcatcc	9600
ccgctaccog	tgaggagcagg	tcagggactc	gcgtgtctgt	ggcacaccag	gcctgtgaca	9660
ggcggtgttc	catgtactgt	ctcagcagtg	gttttcttga	gacagggctc	cgctcgctca	9720
cccaggcgag	agtgcagtg	cgcaatcacg	gctcgctgta	gcctcaatct	ccctgggctc	9780
aggtgatcct	cctgcctcac	cctctgagta	gctgggacta	cagacacata	ccaccacacc	9840
cagctagttt	ttgtgtattt	tttgtggggg	gagatggggg	ttcgctgtgg	tgcccaagct	9900
gatctcaaac	tcctgaggca	caagcgatcc	acctgcctcg	gcctcccaa	gtgctgggat	9960
gacaggcatc	agccgtcaca	cgcagctcaa	tgattttatt	gtggtaaaa	aaacatagca	10020
caaaattgat	gattttaacc	attttaaagt	gaacagttca	ggctgggctg	ggtggcttat	10080
gcttgtaatc	ccagtacttt	gagaggctga	ggtgggcaga	tcacctgagg	tcaggagttt	10140
gagaccagcc	tgcccaacat	gatgaaatcc	agtctctact	aaaaatacaa	aaattagccg	10200
ggcatggtgg	caggtgcctg	taatcccagc	tactcgggag	gctgaggcag	gagaatcgct	10260
tgagcccggg	aggtggagg	tgcatgtatc	tgagatcatg	ccactgcact	ccaatctgtg	10320
tgacagagca	agactctgtc	ttgaaaaata	aataaataaa	aaaaatttta	aaaagtgaac	10380
aattcagggc	athtagtatg	aggacaatgt	ggtgcaggta	tctctgtctac	tatctacttc	10440
tagaacactt	tcttctgccc	tgaaggaaac	cccatgcccc	ccggcactca	cgcccattct	10500
ccctctctct	ccagcctctg	tcaaccacta	atctactttc	tgtctctggg	ggttcacttc	10560
ttctggacgt	tttgtgtgac	tggaatcctg	caatatgtgg	tccttgcgtg	tggttctttt	10620
ccatagcatt	gtgttttcca	gattcaccca	cacattgtcg	cacgttatca	gaatctcatt	10680
cctgactggg	tgccagtggt	taggcctgta	atcctaacat	tctgggaggc	caaggcggga	10740
cgatcacttg	agcaggagt	ttgagaccag	cctggccagc	ctagcaagac	cccagctacc	10800
aaaaaatttt	aaaagttaac	tgaacgtggg	ggtgggtggg	acttgtggtt	cccagctacc	10860
tgaggaggctg	aggtgggagg	atcgcttaag	ccaggagggt	caaggctgca	gtgagctatg	10920
atcgaccac	tgactccag	cctggacaac	agagcaagac	cctgtctgaa	aaaaaaaaca	10980
aaaaaaaaag	ttcctttctt	tttgtggctg	gatgacatcc	cattgtatgg	ccacagcaca	11040
ttttgtttgt	ctgtttatcg	ggtgggtggg	agtgggtttc	accttttgtc	tcctgtgaat	11100
aatgctgctg	tgaacatttg	aattcaagtt	tttgtttgaa	cacctgttgt	gaattatttg	11160
gatatatgtg	taggggtagg	attgctgagt	cctatggtaa	tgtaggtttt	gacttactga	11220
ggaaccatta	aactgttttc	aacagtggct	gcgcggttct	gcatccccac	cggcagtgtg	11280
tgagggttct	gactttacct	cctcacaaac	gcttcttttc	catttaaaaa	aatattcagc	11340
caggtgctct	ggctcacgcc	tgtaatccca	gcactttggg	aggccgtggc	gggcggatca	11400
cctgagggtca	ggagtctcag	acgagcctgg	ccaacatggt	gtaaccccat	ctctaccaa	11460
aatataaaaa	ttagccgggt	gtggcagcgg	gcgcctgtaa	tcccagctac	ttgggaggct	11520
gaggcaggag	aatcacttga	acccgggagg	cagagggttc	agtgagccaa	gatcgcgcca	11580
ctacactcca	gcctgggtga	caagagtga	actccatcta	aaataaaaaca	aaaataaaaa	11640
taaataaaaa	tttattaaaa	cattcatcac	agccagccta	gtgggtgtcc	catgtggctt	11700
tgctctgcat	ttccctgata	actaggatgc	tgagcgtctt	gtcccaggct	tgccacacct	11760
cagcactttg	agatacgtcg	cacagtcccc	atttgcgaa	gagaaatgag	gtttagggaa	11820
cagcagctgt	gtcatgtcac	acagcagaca	gggggtctct	gagccgtctg	acccacagc	11880
cgaccaagct	ccaatcctta	ccgcctccta	gtgttggtga	tgtagccag	ggtgctcca	11940
catttttccag	atgagaacac	cgaagctcaa	aacaggagcg	ttttgtccac	attggatata	12000
cgatgtctgt	ggtttggtcc	tgaagtcact	ttatatctca	gtggtccaga	ctggagttag	12060
acagggggtt	ctggggaatg	gggaagggtg	ctcagggtgaa	aggaagggaat	tccagattct	12120
ccatactgtc	cttgggaagt	tagaagactc	agagggctctg	gcaaagttag	acaaagcaag	12180
agaaatgcag	tcaggaggaa	gcggagctgt	ccagggaacag	gggggtcgca	ggagctcacc	12240
cccaggaaact	acacttgctg	gggccttcgt	gtcacaatga	cgtgagcact	gcgtgttgat	12300
taccactttt	tttttttttt	ttgagggtgga	gtctcgctct	cttgcccagt	ctggagtgca	12360
gtggcacgat	ctcggtcac	tgcaagctct	gcctcccggg	ttcatgccat	tctcctgect	12420
cagcctcccg	cgtagctggg	actacaggcg	cctgccaccg	cgcccggcta	atthttgtat	12480
ttttagtaga	gatgggattt	cactacatta	gccaggatgg	tctcgatctc	ctgacctcat	12540
gatccgccc	tctcggcctc	ccaaagtgt	gggattacag	gcgtgagcca	ccgcgcccgg	12600
cccgaatttc	cactttaaga	atctgtctgt	acatcctcaa	agccctatac	acagtgtctg	12660
gttgctatag	ggaatatgag	gcttacaggc	catggtgctg	gacacacaga	agggacggag	12720
gtcaggagggt	agaaggcg	agagaggga	caggcggagg	tcacatcctt	ggctttcaaa	12780

atggggccagg	gagagacacc	ctctgagcat	ggtaggacag	gaaagcaaga	ttggaacaca	12840
ttgagagcaa	ccgaggtggc	tgggcgtggt	ggcttacgcc	tgtaatccca	acacttttgg	12900
aagctgaggt	gggtggattg	cttgaggcca	ggagtccaag	accagcctgg	ccaacatggg	12960
gagaccccg	ctctactaaa	tatacaaaaa	ttagccaggc	gtgatgggtg	atacctgtaa	13020
tcccagctgc	ttgggagggt	gaggcaggag	aattgcttaa	acctgggagg	cggaggttgc	13080
agtgagccga	gatcccgcca	ctgcactcca	gcctgggcca	cagagtgaga	ctccatctca	13140
aaaaaaaaaa	aaaaaaaaaga	taaaaagacc	aaccgaggaa	ttgaagtggg	ggggcgtcac	13200
agtagcagaa	gggggatcgt	ggagcaggcc	accctgtggg	catgcactgg	aagctcatta	13260
cctgacgatt	tggagctcat	cactgggggc	ctaaggagaa	tagatactga	aggatgagga	13320
gtgatggcgc	ggggcacggg	tgtctttggt	ggccagaact	tggggactgc	tgggggtgct	13380
cactgcaggc	cttctcagcg	ccctttatat	gcttacacag	gctgtttcta	agaggggggat	13440
acattgcata	agcgttttca	gactacctca	tcatgggtcc	ctttctttac	cctctgtggc	13500
cctggtggcg	cactctctgg	gaaggtgcag	gtggatgccc	agacccgccc	tgccatccac	13560
ctgcacgtcc	agagctgact	tagcctcgag	attgctgctg	gcacctcctg	ccccgggaca	13620
cctcggtatg	gcccgtggag	atgctggctc	tgtgttttct	gctggagttt	ggtgctgctt	13680
ttcctcctgc	aagtggccac	cgctcttggg	tatgtcctca	ggcttctgcg	agtcattggc	13740
gcttctcagg	tccttgccca	gcgccaggag	caaaccctcc	tggcactttg	ttcaggggtg	13800
gatgcgccag	tgttctctgt	gtggaccccc	atctcacatg	agggctcttg	gcctgcaggc	13860
tcgttcagga	aacaccgcgt	gagtacgcag	tgtgtgccag	ctgtgtccca	ggcaatggcg	13920
gggacagtgg	ctgctgctgg	ggttgtggtg	gcttctgggg	actctgggga	cagctgaggt	13980
gcaaggagcc	acgggtcctt	gaggtgcag	ttggactcca	ggtggaaggg	atgggttggg	14040
gaggtataaaa	tggggtcagg	gaggagacac	atttggaaaca	atgggaacat	ttttaagatg	14100
ctatgtcggg	aggcaacaag	gtggccaacc	caggtgctga	ggagcccaca	ccagccctgg	14160
acgtgttttg	ccgtcacctt	ttgctgggga	gtggtgggag	agaggattcc	gttccacgtg	14220
gtggtgtgcg	cagctgggct	gtgtggagct	gggcgctagg	aggaagggtg	tttctgcggg	14280
gctagccggg	ctctgccttt	gaacacaatc	aggctccagg	ttttcagcat	ccagtgcatt	14340
agaggacttc	acgggcagct	gtggctgac	ccttgatgaa	ttgggagaag	aacaaagggt	14400
tatgaaatga	ggtttcatgt	agatggcatt	agagacgccc	acaacagatt	tacagagtgg	14460
agcggagacg	gcggatgggt	ctgggaggcc	cctcctgctg	gccttgactg	tgacagctgt	14520
cctgggaatc	agcttccagg	ccgcccacgc	agcctgactg	acacacacag	gggttttagc	14580
cccattcctg	gaccagctgt	tgccatcatc	agtgcacagt	gggagtggcg	gtggttccag	14640
ccctgggcac	cctccccacc	tgctggggcc	cacccagggc	agtcctgaca	cctacaggtt	14700
gcttggagcc	gcattccgag	cctgccccac	cacgtgtgaa	gcccagagtg	tcgtgggctg	14760
aggtcccctg	attgcatccc	cacttccctt	ctgcttcaca	tagctgcctc	ttctcaccgt	14820
ttttccagcc	tcctgggcta	ggaattccag	tgttgtgctg	gctttgcccc	aggacacctc	14880
cttagccctc	ttcctgagtc	tagagccccg	ggggttggaa	gttctggccc	ctgggacacc	14940
tgcagccaca	ctcagcttct	cctgtgagcc	tccagcatgt	cccctcagga	ccaagccctc	15000
acgttcttgc	ctccccgccc	acctgggctc	agccagggga	aggcctggct	gggagcgtct	15060
cccctctgcc	ctgcccttct	cccctctacc	ctgcccttct	ctcctctgcc	ccgccatggc	15120
ttttatatcc	tgtgccacaa	gacatggctg	tgtgtgaaag	tggcagggtc	tggcatctct	15180
gtgggtctct	gaggcccacg	ctccagtgcc	actcttccca	cccgttgccc	gtgccctcat	15240
gctggaggga	cagcccagcc	ctctcccgaa	ccccagcccc	atgtgccccg	ctgcccccg	15300
ccctctcccc	tgggaagccg	ggtcactcca	gccgtatgcc	atggtgggga	catcctgctt	15360
ccttggcctt	ccagggaagg	tcctctttcc	aaatggcgac	acctgggtccc	tgccctggagg	15420
ctggaagctg	tggcccttgt	atgcccctcc	agggctctgt	cgctcggttg	gcccaggttc	15480
ccatcaccgt	catcatcacc	atcatcattg	tcatttctgt	tgtctgtgag	ccggcctggg	15540
ctcccagagc	agagaccctc	tgaggtccag	cctgagttgg	ggtctccgtg	ctgacccctg	15600
acggggactc	aggacgtacc	aggtctgggt	caggagtgc	ccccaaacct	cgtgcccttt	15660
gacaggcacc	cctgactttt	gctaagtggg	tggaggtgac	atcacttaca	gcgggagtga	15720
tgggacaggg	tctgttggct	gcactgtgct	cccagggatc	tggggagagg	ctatatccct	15780
gggctttggc	actgcagagc	tgtgtgtgtt	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	15840
tgtgtgtgtg	tgtgtgtgtg	tttgctgctg	cgcacatgtg	tataagatct	ttttttatta	15900
catgaagcaa	gataactgtt	gctgtttcct	tttgggtttt	gtgttcaaca	gagtgggggt	15960
cttcttccct	cagacaacag	aactctcccc	tttaaacacg	tgctgtcaga	gggtgggtct	16020

tgggctcatg	tctgttttga	cagccgagtc	agaggaaaca	caggggttctt	cataaaaaaca	16080
ctgcacagca	ggcgactgtc	cagagtcage	ctgcaggacg	gcagcagccc	tgccccctcag	16140
agcacagcta	gggtgggctg	ctttgggata	tcccgtcatt	ccctcccage	tggcagccgg	16200
cggccggccc	attccttggg	gtgctgggtca	ggggggcggtg	cgctgctct	gctcaccctg	16260
ggaatgggac	agaagctggc	agctcggaga	ggacagggct	ggacccttgg	gtggcctctg	16320
gctggacat	ctcattgtcc	tcagacacag	cctctcgggt	ctagtttcat	ttcctgaaaa	16380
acaagtgcac	agaactagag	caggagtcga	gagctacggc	ccccggggcca	gatccagccc	16440
tgccacctgt	tttcacacca	tgctcaagct	gagtggtttt	tacatttttt	aattacttga	16500
aaaaaaaaa	gccaaaggag	gtttcatgac	ccatgaaaat	tatatggaat	tcaaaaaaaa	16560
aaaattatat	ggaattcaaa	tttcagtgtc	cataaataat	ttcttgagac	aggggtctcgc	16620
tctgtcacc	aggtctggag	gcagtgtctat	ggcatggctc	gctgtaccct	tgacctccca	16680
ggctcaagcg	atcctcctgt	ctcagcctcc	tgagtagctg	ggactacggg	tgtgtgccac	16740
caagcccgcc	taattttttt	tttaatttttag	taaagacagg	gtctttctat	gttgccagg	16800
cttttctgga	actccatctt	ggcctcccaa	agtgtctggga	ttacaggctc	gagccacgga	16860
gccagcctg	tttttgtttt	ttcactgata	aagttttgcc	gggtgtggta	gtgtgtgcct	16920
ctagcgattt	gggaggctga	ggtgggagga	tcgcttaagc	ccaggagttt	gaggctgggc	16980
tcaagtgate	aggaggtgaa	ctatgatcat	gtcattgcat	tccagcctgg	gtgacagagc	17040
aagaacttat	ctcttaaaaa	tatatattta	aaaagtattg	gggtgtggtg	ctcacgcctg	17100
tgggtcccagc	tacttaggca	tctgaggtgg	gaggatggct	tgagccagg	agtttgagg	17160
tgacgcgagc	caagatcgtg	tcactacact	ctagcctggg	tgacagagcc	cagacctgc	17220
ctctttaaaa	aaaaaaacca	aaaaacatgt	attggaacac	agccatgcct	gttcagtcac	17280
gtgctctcca	tgctgctttc	tgctccagag	acccttatgg	cctgaaagct	gaaaaatattt	17340
tctatccttt	acaaaaaagt	ttgctgacct	ctgtcctgga	aaattcatct	cccaagttct	17400
cttccggcac	tggcgcttct	gggtgtccta	aatttggccc	ctgttatctt	tgaactctgt	17460
tttggctctg	ttccctccca	ggagccagga	caggcacgtt	ctctgcatct	tgtcccctga	17520
cggccagagg	cttggctcgg	ctcaggcatt	cttggaaata	tctggctcca	ggaaaggcag	17580
aggcctcctg	agtcagccca	gagggaaacct	gccccaggct	tgggggaggc	ctgaccagc	17640
agagtggctt	ttgccgatgg	gttgggccc	tcaagatgtg	ctgaaagtgt	tcctcagaag	17700
gccacttttg	gattccttcc	tccagtatta	gagcaactga	gagctgctca	ttgcaagcct	17760
gatgttttcc	cagttggccg	ggtccaccgg	gtgcctggg	attctgggat	ctgggtggaa	17820
agtagggggc	ttggggggagt	gtcctgggtt	ctggaatcca	gggtggcaagt	ggtgagggttc	17880
agggagtggc	ttctgagcca	ccataggggt	ctctgtggga	ggctctgccc	atccaggaga	17940
ttccgcaggc	cctgccggcc	cagagccagc	gtcttgcgct	tgccgaggct	acagccagcc	18000
ccagccgggt	ggaacagccc	gtgcctcct	ctcactttgt	tttggggcca	cctgggagt	18060
tggagcaagg	gtagagaggg	aggaagtggc	tgccggccgc	tgcccagcac	ccttgtttgc	18120
cttgggcccc	ctgtgggctc	ctttttattg	ctcttcaatg	aagccaggga	aatggacttc	18180
cttgctcac	ttcagttcaa	catgtctgga	agtttggtat	taaaattaa	aaagtgtgga	18240
aatagagcaa	gaagagaaaa	atctctccaa	gagataatag	tgacctctga	gctgggcgcg	18300
gtggctcacg	cctgtaaatc	ccagtacttt	gggaggctga	ggcgggcaga	tcacctgagg	18360
tcgggagttt	gtgaccggcc	tgaccaagat	ggagaaaccc	cgtctctact	aaaaataaat	18420
aaataaataa	ataaataaat	acaaaattag	ccaggcatgg	tggcgccctgc	ctataatccc	18480
agctaaggca	ggagaatcgc	ttgaacctgg	gaggcaagg	ttgcagttag	ccaagatcac	18540
gccattgcac	tctagtctgg	gcaacaagag	tgaaactccg	tctcaaaaaa	aataaataaa	18600
taaaaaataa	aaatagttag	ctctggccag	gtgtggcagc	tcataaccgt	aatcccagca	18660
ctttggaagg	aaggccgaga	tgggcagatt	gcttttagcac	aggagttaga	gaccagcctg	18720
gccaacatgg	tggaaaccca	tctctacaaa	aatagaataa	aatttaagag	gtaatagtga	18780
ccttttggtg	gatcgaaacc	tggattgctt	tctttttcta	aatgctgatt	cttttctttg	18840
tgggtgtttg	gttctgtgcc	gatgtccctc	ccccagccct	gttattgtga	gtggaagaag	18900
gggaaagggt	tcgcccgcga	ctgtgagccc	ctcctctcac	gctgggtgtc	cttggaagaag	18960
cctgcacttc	ttcattgtac	gccagggtct	gggtccctccc	tggagtgggt	ctgtgctgct	19020
gggatggggc	caacccctca	gatgttttct	gagtgctaca	cacagggtgt	tgcatctcatg	19080
gcctttgcgt	gtcttcctgt	tgtggaggca	aaaatgtgaa	gaaccctaga	tgatttttggg	19140
accagggtct	catcacctgc	tgttcattgc	acaccggagc	atccaggcat	gggtggagag	19200
ctcagacttc	caggcacggg	cgcaggggct	ggtctaacca	tgttcccgc	cgcctgctcg	19260

tcagaaccgc	ctgttgaggag	ctgttatcat	gataccatac	ctggggccctg	ggctatccga	19320
ttctgactta	attgctccag	gttggggcca	ggccgttggt	tgctgttttg	ttgtttcttc	19380
tgtgacgtta	gccactgggc	taatctgagc	ccctcagtta	caggtggaga	aactgagacc	19440
catgggggtg	caaggacttg	ccgaggaccc	agagcccctt	gggggcagag	ctgaggcggg	19500
gcctggcttt	gggtcccaga	gcttccagtc	cccttcccgc	tctcctaaca	gctttttttt	19560
ttgagacaag	atctcaccc	gtcaccagc	ctggagtgc	atggcatgat	ctcggctcac	19620
tgcaatcttc	gctagctgcg	ttccagcgat	tctcctgcct	cagcctcccg	agcagctggg	19680
attacaggtg	tgtgccgcca	tgcccagctc	gttttttttt	gtacttttag	tagagatagg	19740
gtttcaccat	gttggccagg	ctgatctcga	actcctgacc	tcaaagtatc	cgcctgcctc	19800
ggcctcccaa	agtgttagga	ttacaggctg	ggatcacact	gtgcctggcc	ctagcagctt	19860
tgtcctgtgc	catccaacaa	cagatgaccg	aagtctttgt	ttcttaacat	gcattccatc	19920
tgccctacag	ttttgccacc	tgcaaaacag	aggacttgct	gctttttctg	taagctggaa	19980
atgtaatctg	gtagcaggag	gcctgtggaa	gcttgccctt	aatggccttg	tgtctctttc	20040
atcctgtcct	gagagccgga	gaacttggtg	gttgcaccta	actcaacctt	cctgttaaca	20100
tacagttctg	caggctcatg	gatcatcaga	accacgtcct	atctcacgcg	gctgtatgct	20160
tccgttggtt	cagggtgttt	taccttgaca	gtattttctc	ctcgggtggc	tttgcggtgg	20220
ttgcttttaa	tcagcattga	ctcttcaaga	aaaatattta	gctgctacat	ctcagaggag	20280
acagggtgga	aagcatctga	gacctgcagg	ctcagactta	gaaccagaag	tgccctcaga	20340
gttcatccgg	ccctgaccca	gcgggaaatg	agttcacaga	gaagcgggag	aactttgccc	20400
caggccctgc	cggtgtctat	aactgcccc	ggctccttaca	tttgctccag	gtcctgcccc	20460
aggccctgca	gttgctcata	actgccccag	gtccttata	ttgctccagg	tcctgcccc	20520
ggctcctgag	ttgctctgtg	tggtgggtgt	gatctggagc	cctccgcccc	ttgctgcacc	20580
tggggcaggc	attgctaatt	gatcccagg	ctccttctct	cggagcacgc	cctgggtctc	20640
caggcagccg	ctgcctgtca	gcctgcagtg	gttcggggaga	ggacacctgc	ttgcctgggtc	20700
tgttccaaat	cttgcttctc	atcccagcac	aggtaggggg	tgctatggga	aagggtacct	20760
cagttggccc	tgtcactgct	ctatcagctg	gggacgtggc	atcctagtga	aaacatcatg	20820
gccgggcgcg	gtggctcacg	cctggaatcc	cagcactttg	ggaggctgag	gagggtggat	20880
cacttgaggt	cagaagtctg	agaccagcct	ggtcaacatg	gtgaaaccca	tctctactaa	20940
aaatacaaaa	attcgccagg	tgtgggtggc	ggtacctgta	atccgagcta	ctcgggaggc	21000
tgaggcagga	gaatcgcttg	aacctgggag	gtggagcttg	cagtgaagcc	agatcttgcc	21060
actgcactcc	agcctgggca	acagagtga	acgctgtctc	aaaatctcaa	acaaacaaac	21120
aaacaaaaaa	caaacaaaca	aagcgtcatt	tatccagcac	ccctggggaa	ccatgctacc	21180
tggtgtttta	tggtacctgg	caaggtgcag	gtgaagttgc	tgctcttggg	cattgaaccc	21240
gtcttggttg	gggcagctca	ggccccaggc	agggtccggg	ttggctctcg	ttgggtgtggc	21300
cctggcccat	ccagacctat	atttctgccc	tcctgcaggt	gatcaatgtt	gatgggacga	21360
agaggcggac	cctcctggag	gacaagctcc	cgcacatttt	cgggttcacg	ctgctggggg	21420
acttcatcta	ctggactgac	tggcagcgcc	gcagcatcga	gcgggtgcac	aagggtcaagg	21480
ccagccggga	cgctcatcatt	gaccagctgc	ccgacctgat	ggggctcaaa	gctgtgaatg	21540
tggccaaggt	cgctcggtag	tccggggggg	cccaagccat	ggctcagcca	tgagacttg	21600
catgaggagg	aagtgcagg	tccatgcctg	ggcataagtg	ttgagctcag	gtgccccgac	21660
ctggggaagg	gcaggacagg	aaaggtgaca	gtatctggcc	aaggacagat	gggaaggggac	21720
caagggaagt	gattagggag	tggttatgga	ctagggaatg	cggtaacaat	ggttagaaag	21780
tgactaacat	ttgttgagca	cctgctgtgt	gcccggccct	ggccggggagc	cttcgtgccc	21840
acagtgaacc	cgctcgcaaa	tgtagtctct	tgccctactc	gcactgggga	gcaggacgca	21900
gagccgtgca	tctcacaggt	gccaagctca	ggactccctc	ctgggtctgc	ctgggctggg	21960
ctgtgcttgt	tgccccgtg	gcccacgcat	gtgcaccttc	cacctgaaag	ccaggatctt	22020
caggacgctc	cccaggagg	tcgttgtctg	gcacaatgat	ttgtctcttc	ctgaaaagg	22080
gacagagtta	cactggagag	agcagcatcc	aggtgcggca	gggacaggcc	tggggctcgc	22140
gggcaggggac	tctgtgtcct	gcccgggtcc	cacactgcac	ctgcttgctc	gaggcactca	22200
gtcaatcttt	gctgatgaag	gatgagagga	cagaggacgt	gatgcttgct	gctgcattgc	22260
ctgcagtcct	gggtgagatg	cccgggttga	ctctgctgcc	cgtcgggtgg	atgtgatgtc	22320
agatcccccg	ctttaaaata	cgagggagct	gggaattgag	ggagcagggt	ggggcagaaa	22380
gcacagcccc	gtggaagcct	ggagctgagg	cagtgtgggc	gacccctgga	gcagtgaagt	22440
cttccttcat	ggccttcatc	gcacctgca	gtcctcatgt	aggggatgcc	atccatgaat	22500

ttagttttcc	cagcctcctt	taaaaacgcg	ttcatgctgg	ggccggggca	gtgcagtggc	22560
tcacatctga	aatcccacca	ctttgggagg	ccgaggcggg	tggatcatga	ggtcaggaga	22620
tcgagaccat	cctggctaac	aaggtgaaac	cccgtctcta	ctaaaaatac	aaaaaattag	22680
ccgggtgcgg	tggcggggcg	ctgtagtccc	agctactcgg	gaggctgagg	caggagaatg	22740
gcgtgaaccc	gggaagcgga	gcttgcaagt	agccgagatt	gcgccactgc	agtccgcagt	22800
ccggcctggg	cgacagagcg	agactccgtc	tcaaaaaaaa	aaaaaaaaagt	acaaaaaaa	22860
aaaaattagt	ctgggtgtgg	tatcacgcgc	ctataatctc	actactcgag	aggctgaggc	22920
ggagaattgc	ttgaaccag	gaggtagagg	ttgtagttag	cccgatcgt	accactgcc	22980
tccacctggg	caatagagcg	agactctgtc	tcaaaaagaa	aaaaaaaaaa	agaacattta	23040
tgccaggtgt	ggtggctcat	gcctgaaatc	ccagaacttt	ggaagactga	ggcaggagga	23100
tcacttgagc	ccagaaattt	gagagtgtct	tccctgggca	acatagagag	acctcatctc	23160
taccagaaaa	aaaaaaatta	gcccggcatg	gtggcatatc	cctgtggtcc	cagctactta	23220
gggggctgac	gtggcaggat	cacctgagtc	tggaggcaga	ggttgaagt	agctgagatc	23280
atgccactgc	actccagcct	gggtgacaga	cagagaccct	gtctcaaaaa	aaaaaaaaaa	23340
aaaaagcatt	tactatccac	catggaaggt	gagactgacc	tgtgagtgat	tgttcaaaga	23400
acaaaaaata	aaccccagag	ataagacaaa	aggggtgcctc	catgggggtg	tgatttaaag	23460
ctgagaaatt	gggcttcttc	cccctcccct	ctcaccccg	ggtttgctaa	aggagatggg	23520
aaaaaggatt	cttttttttg	ctgaaatatt	taacactaaa	ttaaagccaa	ttttaacagc	23580
actttggttg	atgagtga	ttaacagact	ggccaaaaat	aaacgaacgg	tctgtactat	23640
gtgaaaaaga	ggcagctttg	gccatgctgg	gccaatgtga	gttttcaggg	ttgctgggaa	23700
tgtctgtgaa	tcggaggaag	ggcctagctg	ggactctcag	gagccaaggc	cctgaggggc	23760
aacttgccctg	gtccctgccc	tgaggcgctt	actgctttct	tcctggggca	gatcacaggc	23820
ccggaggctg	gacctgagg	ctggcactct	tgccgagctg	ctccctgact	tcctgacct	23880
gctcctttca	gcagccttgc	tgcactttag	tttccttgaa	tgaaaaatgg	ggatgagaat	23940
agctcctacc	tccaaggtga	atggagttag	ttcggacagg	tgactccctg	ggaccagtgc	24000
ctggcgccctg	acaaggtcca	gtcagagccc	gcactgctgt	tactgatacc	cttggctgta	24060
ccaggggaga	acttggttgc	cattgccagg	tgcttctcca	ccacccccac	tactgtccct	24120
gtttgatgtg	tggcgggaat	aaagctgtgc	acattggagc	ttttggcaca	tcctggcctt	24180
caggtgaaag	gtgctgtgtg	gtttgagggt	ttagcctggc	caacccagcc	atgaggtcgg	24240
acctgacctg	ggggtgagtc	ctgagctcgg	caccctgag	ctgtgtggct	cacggcagca	24300
ttcattgtgt	ggcttgcccg	cacccttttc	cctgctgggc	tggtgatgtt	tagactggag	24360
cctctgtgtt	cgcttccagg	aaccaaccgg	tgtgcggaca	ggaacggggg	gtgcagccac	24420
ctgtgcttct	tcacacccca	cgcaaccggg	tgtggctgcc	ccatcggcct	ggagctgctg	24480
agtacatga	agacctgcat	cgtgcctgag	gccttcttgg	tcttcaccag	cagagccgcc	24540
atccacagga	tctccctcga	gaccaataac	aacgcagtgg	ccatcccgt	cacgggcgtc	24600
aaggaggcct	cagccctgga	ctttgatgtg	tccaacaacc	acatctactg	gacagacgtc	24660
agcctgaag	tagctggggc	cagaacgtgc	acacaggcag	cctttatggg	aaaaccttgc	24720
ctctgttcct	gcctcaaagg	cttcagacac	ttttcttaaa	gcactatcgt	atttattgta	24780
acgcagttca	agctaataca	atatgagcaa	gcctatttaa	aaaaaaaaaa	gatgattata	24840
atgagcaagt	ccggtagaca	cacataaggg	cttttgtgaa	atgcttgtgt	gaatgtgaaa	24900
tatttgttgt	ccgttgagct	tgacttcaga	caccccaacc	actcccttgt	cgggtgccgt	24960
ttgctcagca	gactctttct	tcatttatag	tgcaaatgta	aacatccagg	acaaatacag	25020
gaagactttt	tttttttttt	tttgagacag	agtcttactc	tggtgcccag	gctggagtac	25080
cgtagcgtga	gctcagctca	ctgcaacctc	cgcctcccag	gttcaagcga	ttcttctgcc	25140
tcagcctcct	gagtagctgg	gactacagac	atgcaccacc	acacccagct	aatttttttt	25200
atatttttag	tagagacagg	gtttcatcat	gttggccagg	ctggtcctga	actcctgacc	25260
tcaggtgatc	tgccgcctc	ggcctcccaa	agtgtgaga	taacagggtg	gagccaccgt	25320
tcccggcata	ggaaaacttt	ttgccttcta	aagaagagtt	tagcaaaacta	gtctgtgggc	25380
tggccttctg	attctgtaaa	gaaagtttga	ttggtggctg	ggtgcgggtg	ctcacacctg	25440
taatcccatc	actttgggag	gccgacgtgg	gcataatcacc	tgatgtcggg	acttcgagac	25500
cagcctcacc	aacgtggaga	aaccccgctc	ctactaaaaa	tacaaaaaaa	aaattaaccg	25560
ggcatggcgg	cgctgcctg	taatcgagc	tactcaggag	gctgaagcag	gagaattgct	25620
tgaacctggg	aggcggaggt	tgtggtgagc	tgagatggca	ccattgcact	ccagcctggg	25680
caacaaaagt	gaaactccgt	ctcagaaaaa	aaaaagtttg	attggtgtaa	ccaaagcgca	25740

tttgtttatg	gattgtctgt	ggcagctttt	gttctgccga	gatgagttgt	gacagatctg	25800
tatgggctct	aaagcctaaa	acatgtgcc	tccgcccctt	tacagaaaaa	gtgtgctgac	25860
ctctgttcta	aagtattgga	caactacaat	gtttgctcat	ttattattct	atgatttggt	25920
ttctgctttt	tgttggtgtt	gttggtgttg	agatagggtt	tccctctgtc	actcaggctg	25980
gagtgcagtg	gtgtaatctc	agctcactgc	agcctcgacc	tcctgggctc	tagtgatcct	26040
ctcatctcag	cctccctagt	agctgggact	acaggcacac	accaccactc	ctggctgatt	26100
tttttttttt	tttttttttt	ttgtggagac	agggtttccg	catgttgccc	aggctgggtt	26160
caaactccta	ggctcaaaca	cccacctcag	cctcccaaag	tgctgggatt	acaggcgtga	26220
gccaccatgc	ccagcctatt	ctactgtttg	tattacatag	ctttaaaaga	ttttttatga	26280
ctttaagtca	caagggttct	ttgtagaaaa	aaatatatat	ataggaaagt	ataaaaaaga	26340
agtaaaaaat	gtccataacc	tctccagcca	gagacgaccg	ttgctgacac	ctcagcatat	26400
tgcccttaag	tcttttttct	ctaagatagc	atctctcttc	atcacagtca	tatgctacgc	26460
agaattctgt	atcctgattt	tttcaactga	cattacaaca	ggtatttgat	ggcgtgtgta	26520
caaactcttt	ggcacaaatc	tttaaagtga	tgaaatactc	cactgcacag	atgtttgctt	26580
ttaggcttaa	ctgttctttt	attttgctgt	tgctgggttac	agccgggcac	agtggctcat	26640
gcctgtaatc	acaacacttt	gagaggggtga	ggcaggaggga	tcacttgagc	ccagaagtgt	26700
gagaccggcc	tgggcaacat	agtgaagccc	catctctaca	aaaaactttt	ttaataagtc	26760
gggcgtagt	gtgcatagct	gtagtcccag	ccaccaagga	ggctgagttg	ggaggattgc	26820
ttgagcccca	ggaggttgat	gctgcagtga	cctgagatta	ctccactgta	ctccaacctg	26880
agcgacagag	caagacttgt	ctggggaaaa	aaaaaaaaaa	aatatatata	tatatatata	26940
tatatataca	tatatacata	cacgcacaca	cacataatat	aaaaatata	atattataat	27000
atataatata	taataataaa	atataatatt	ataaataaaa	tttataaatt	atatttataa	27060
gtaaatatat	aatatataat	ataaaaatat	atattatata	atatataata	aaatatataa	27120
tataaaaaata	tatatattata	aataatatat	aatacatact	tataagtata	tatttataat	27180
atatgtaatg	tatatttttt	aatgtatgat	atataatata	catttataaa	tacacattta	27240
tattattttta	tataaaatat	atataaaatc	tccaagttgc	tttttccaaa	aagggtgtctt	27300
gctgcatttc	aaacattcat	ttaaaaactt	gaatgctggg	gatctgggtc	agaatgtgtt	27360
cagtagctgc	tgccagtggc	caagcatctc	gggagatgtc	tacaaaacac	gctgggtctg	27420
gcctggcggtg	gtgggtcacg	cctgtaatct	cagcactttg	ggaggctgag	gcagggtggat	27480
caactgaggt	ctggatttctg	agaccagcct	tgccagcttg	gtgaaacccc	atctctacta	27540
agaatacaaaa	aaaattagcc	aggcgtgggtg	gcatgtgcct	gtaatcccac	ctacttggga	27600
ggctaaggct	ggagaatcgc	ttgaacccag	ggggcagagg	ttgcagtgag	ccgagatcgc	27660
accattgcac	tccaggtctg	gcaagaagag	cgaaactccg	tctcaaaaaa	aaaaaaaaag	27720
atgctgggtc	ctaaaatgtg	gccctttttc	tctcacctg	ctgccagacc	atcagcccg	27780
ccttcatgaa	cgggagctcg	gtggagcacg	tggtggaggt	tggccttgac	taccccgagg	27840
gcatggccgt	tgactggatg	ggcaagaacc	tctactgggc	cgacactggg	accaacagaa	27900
togaagtggc	goggctggac	gggcagttcc	ggcaagtcc	cgtgtggagg	gacttgggaca	27960
accogaggtc	gctggccctg	gatcccacca	aggggttaagt	gtttgcctgt	cccgtgctgc	28020
cttgtgttca	cctcgatatga	gacagtgcgg	gggtgccaac	tgggcaagggt	ggcaggctgt	28080
ccgtgtggcc	ctcagtgtat	agagctgtac	tgatgtcatt	agccttgatg	gtggccagga	28140
ctggtagggc	cctcagaggt	catggagttc	cttcgtggag	cgggtgctga	ggctgtatca	28200
ggcacagtgc	tggctgcttt	cacctgggccc	gtctcaccca	agtgtccatg	gagcctgcgt	28260
aggggtgggt	tctgtgtcga	ttttacagat	gcagaaacag	gctcagagaa	accgagtgc	28320
ttccctaagg	tcacataccc	agtttagagca	gagctgggccc	aggaagtgtc	gtctcaggct	28380
cctgaccagg	tctccttgct	ttgcactctt	gccaaaacca	tgatccagaa	ctgactttga	28440
gggtcccggga	cctcaggctc	ctccgaaatg	gctccttgga	ggctgctgag	ccacagctta	28500
ggacccacct	cgagaggcaa	atgtgctttg	agctgccagg	cgtcctgggg	gccctgcctt	28560
gggcacgggg	ttcagacagg	ccccagatgt	gtggggcgctc	tttctggact	tgagttttct	28620
tttctgtgtg	gtggacacag	tgctcacccc	ttaaagcacc	tgtgatgtgt	gcagcagccc	28680
aatccctgcc	tgctgcctgt	tctgctaggg	aaggaaggaa	gacttcagga	tggcaggaca	28740
acagaaaagag	gtccagggtt	tagagcaagg	gcagggtcaaa	cttagaaaa	tctggaatga	28800
ggatgtgcat	ttcctcttct	ggatctgcta	aaagaagagg	gaaggagggg	ctgctggggg	28860
aggagcccag	agccgagttt	acatccggat	cccgcagggc	ctccctgcc	ctgaggtctt	28920
gttttgtgat	gtgcttgtgt	ccatccctggt	ttctgccgtg	tccccaacat	ccggccaagc	28980

ttaggtggat	gttccagcac	acactcaccc	tgtctgtgca	cctgtttttg	tgtccgtaag	29040
tgggtattta	ctcaccttac	gagtgaacca	ctgtgggaat	tcagggaggt	ggcgcagtga	29100
ccacccctgg	agggatatgt	gtgtggcagg	ggtcgaaggt	ctcgcccttc	cctgcttcct	29160
gcgcgtggct	ttctccagga	cggggagggc	tgagctgaag	aggtggggac	agttgcgtcc	29220
ccccgccacc	cactgtcctg	cggtagagag	agactcactg	agcctgccct	tctcccttgt	29280
gccttccagc	tacatctact	ggaccgagtg	gggcggcaag	ccgaggatcg	tgccggcctt	29340
catggacggg	accaactgca	tgacgctggt	ggacaaggtg	ggccggggcca	acgacctcac	29400
cattgactac	gctgaccagc	gcctctactg	gaccgacctg	gacaccaaca	tgatcgagtc	29460
gtccaacatg	ctgggtgagg	gccgggctgg	ggccttctgg	tcattggaggg	cggggcagcc	29520
gggcgttggc	cacctcccag	cctcgccgca	cgtaccctgt	ggcctgcaag	ttccccaacc	29580
tggcaggagc	tgtggccaca	cccacgactg	cccagcagcc	tcaccctctg	ctgtgggagt	29640
tgtccccgtc	cacccctggg	tgcccttgtc	gcagttaatg	cgggagaggc	tctggtgaca	29700
gctgtttcct	gtgcacctgc	tgggcactag	gtcccagcta	atccctgtgc	caggactcta	29760
atttcacctt	aacacacatg	gtggttttca	ttgctgggga	agctgaggcc	tgagcacatg	29820
acttgccctt	ggtcacatag	ctggtgagtt	caggatcccc	cagagatacc	agggccagca	29880
ctcgatcccc	accagccctt	gaaccccacc	atgtgctggg	attgtgctgg	gagtgtccac	29940
acgcctggga	ccccagggtt	ggtgctctca	tctccttttt	ccagatcatg	agaatgaggc	30000
tcagggaaat	ttgaaaaaaa	cctatcccaa	gtcacacagc	aacaggagca	ggatttgaac	30060
ccagaaaagg	ggaccgcaca	ctctgttctg	ctagagtagt	tagctgtcct	gggtgatatg	30120
gcaggtgaca	ggggcaactg	tgcttaacaa	aggaaccccc	atcccccttg	ccaagttggg	30180
agactagaag	gtcagggggc	gaagctctga	agggccaggt	gcagtggctg	acacctctaa	30240
tcccagcact	ttgtgaggcc	aaggcgggca	gatgatttga	gccaggagt	tcaagatcag	30300
cctgggtaat	gtagtgaagc	gccatctcta	caaaaaaatt	ttttaaaaaa	tagctgggca	30360
tgggtggttca	tgccctgtagt	ccaagctact	tgggaggctc	aggtgggagg	attgcttgag	30420
cccaggagggt	tgaggttgtg	gtgagctgtg	atcatgccac	tgcaactccag	cctgggcaat	30480
agagtgaagc	cgtctccaaa	aaaaaaaaaa	gaagaagaaa	aagaagctct	gaggctccaa	30540
gtccccaggc	acccttgggc	ttgagggcag	acaaggaggg	agagggtcac	ctgggcagcc	30600
ctgacttttg	tcccctggca	aagggacctt	cagtgcctt	ggccctagga	gagcctctga	30660
gcacgtcagc	catgtcgaac	cgctcaggaa	gggcagcaag	aatttggtt	ctgacctctg	30720
cctctcctac	tcgccatctg	cactgggtgt	ggttgtgccc	atcttacaga	tgaggaggct	30780
ggggcatcga	ccagctgaat	gccttgtccc	aggtactgcg	taggcagagc	tggcagttga	30840
accccggtgc	ctggttgtcg	ctgggggtgg	gctgcaccct	gacttgtgag	gccagtagca	30900
aggtttgcac	gtgacttcgt	gaccgtcacc	cagctctgca	gcacatcccg	tgaccagct	30960
catccaggcc	gcatgcaaac	ctgttgccag	gcgagaaaacc	agtcaccgca	cagctgtggt	31020
tgccgtgaaat	gattaagctc	attaatcacc	ccggagttag	gacagactca	gatgaaaacc	31080
agcaaaagcc	ctggaaactc	atgtgaccct	gccaatgagg	gcggccatgt	gcattgcagc	31140
ctggccgtca	ctcctcggtg	cgtgttttgg	acttaaacgc	tccggatgtt	tactgagtgc	31200
ttgattaata	acatggaagg	cctggtctca	ttgctgtggg	agtgaaggat	gcacagccag	31260
gcctgacatg	atgagaacaa	gaacctggag	tctcgctgcc	tgggtggtaa	tccctggccct	31320
gccacttagc	aactgtgtga	ctgtagccag	gtcacttaat	tttgctagat	cctgcctgcg	31380
cttcagtggg	tcttgctggt	tttccaaggt	ggccaaaacac	tttaaggcat	tcattgtggtc	31440
gctaggctgc	agggttgaac	cctggctcac	cccgcagggg	gccgtgtgct	ctgtggcctg	31500
gctgtgcctt	tgctgacacc	gtgcccgtgt	gtgttcatgc	aggtcaggag	cgggtcgctga	31560
ttgccgacga	tctcccgcac	ccgttcggtc	tgacgcagta	cagcgattat	atctactgga	31620
cagactggaa	tctgcacagc	attgagcggg	ccgacaagac	tagcggcccg	aaccgcaccc	31680
tcattccagg	ccacctggac	ttcgtgatgg	acatcctggg	gttccactcc	ttccgcagg	31740
atggcctcaa	tgactgtatg	cacaacaacg	ggcagtgtgg	gcagctgtgc	cttgccatcc	31800
ccggcgggcca	ccgctgcggc	tgcgctcac	actacaccct	ggacccagc	agccgcaact	31860
gcagccgtaa	gtgcctcatg	gtcccccgca	cctcactccc	tcgttagatc	aggctggttc	31920
tgggagctga	cgctgaaagg	agcttctcat	ctgggggtcc	tgggtgtaca	tagatgggtg	31980
ggtaggttgt	gactgcaca	agctgcatga	tgctacctgg	gggtccagg	ccaggctgga	32040
tggacttgtt	gcttcatcag	gacatagata	aatggccaaa	actcctcagc	tggaagggtcc	32100
tgggcaggat	ctttgggtgt	gaaaaccagt	cacaggggaa	gggtgcttgc	tcatactgcc	32160
agcacagtgc	tgagtgcctt	ccatagcgct	cgtttactcc	tcaagcctgg	agggtgggga	32220

gtagcatggt	cccatttcac	gtacaaggaa	cccgatgcac	agagaggtgt	ggcaacccat	32280
ccaaggccat	acaactgggg	tgggttgagc	cggttgtagc	tgtggcaggg	tggctcaaga	32340
gtccctgctc	ctgaaccctt	gccaggcagc	ctggcatcag	ctcggggaat	ttttgccctg	32400
acccttgga	gcaagtgggc	ctctttgttc	tcatgtcagt	gatgagaaga	gtgactttcc	32460
tatggccct	ctggagtaca	ggtgtttcct	gttggcgggc	tcttccccca	tgacatcagc	32520
agcgagctgg	ttatgattcc	ctacgcagaa	cttgatagtt	tataaagctc	tttgtcatcc	32580
aggccccgtt	ggagtctcac	gcagacctgg	tgcaggcgg	ggctggctct	gcctgtccca	32640
gctgcatgga	tggggaactt	gaggcttgca	aagggttaagg	ggctgttcga	ggcccaggct	32700
ggcaggagat	gggcctgggc	cagagtctgg	gacttcccat	gcctgggctg	tctttggtcc	32760
tgttgctcac	catccctccc	tggggccatg	accttagaga	gccaaatgga	ggtgcaggta	32820
accacaggca	aggaggggtt	gccatgactc	agagtccccg	tcctgtggcc	ggcagtaacct	32880
ggtgcaacga	cttggatttc	agaccagcca	ctgtagcccg	ctgacgggtg	gctcgaagtg	32940
ccacagcttc	tgaagccagg	caggactcag	gccaggagac	tctgttagct	ggtgagaggg	33000
agaggccaac	ggatgttctg	gttctgctag	agagctgggt	cttcggatcc	tggtagcagt	33060
gcactgagag	gaggcccagc	ttgattctgg	ggctgccttg	tggtagcatg	tgctgctcac	33120
tgacaccctc	gaggagtgtc	ttctctcggg	cttggtgact	gtgcccgggt	ttccgcagtt	33180
cactggtgca	cacataggca	catagcaaac	cgcacacaca	gtcgtgggta	tgagtttcac	33240
tacattccac	caccagtgtt	cactaccatt	acctgccttc	cgtcttaagt	gttcatcatt	33300
taaaaataaa	tttattgggc	tggacgcggg	ggctcatgac	tgttatccca	gcactttggg	33360
aggctgagcc	gggcctagca	cctgaggtca	ggagttcaag	accagcctgg	ccaatatggt	33420
gaaactccat	ctctactaaa	aatacaaaat	tagctgggca	tggtaggggca	tgccataaat	33480
cccagctact	caggaggctg	aggcaggaga	atggcgtgaa	cccagagggc	agagcttaca	33540
gtgagcccag	atagcaccac	tgcagtccag	cgtgggcaac	agtgcgagac	tccatctcaa	33600
aaaaaaaaata	aataaataaa	agaaaaataa	atztatgata	tatttcaaaa	ataacacatg	33660
tactttgaaa	cagcagagac	acatatgaca	cggagaatga	aattcccat	agcgcacccc	33720
caagagacag	ccctgggtccc	cccgcttttc	ccgtggacct	ccagcggggc	agatgctgag	33780
ccgcctgttg	tgcagtggcg	tgtatcccg	tcctccagct	cctctgtggc	ttacagacac	33840
ccacctgcag	ccctgtcttt	gcctcctcta	gcgcccacca	ccttcttctg	gttcagccag	33900
aaatctgcca	tcagtcggat	gatcccggac	gaccagcaca	gcccggatct	catcctgccc	33960
ctgcatggac	tgaggaacgt	caaagccatc	gactatgacc	cactggacaa	gttcatctac	34020
tgggtggatg	ggcgccagaa	catcaagcga	gccaaaggac	acgggaccca	ggcagggtgcc	34080
ctgtgggaag	ggtgccccgt	gtgcttccca	aggcgctcct	cttgctgggt	tccaggctgc	34140
tgccccgtgc	cttagcagag	ggaggaaaca	gaggatggct	ctgggtgaat	gatgacttgg	34200
gcttcgatta	tgtagtcaca	gggtatgacc	ctgagatgcg	tggaaacccg	agactgtgat	34260
tatatgtaga	aactgggttt	ccccgttgtt	taagtagtca	tggtaggggtc	agaccccaca	34320
ggacttttgt	cttttcaaga	aagaaaatgg	tctgtgtgca	tgcaggggta	gttggtagctg	34380
gttaatccag	gtttatcctt	tattttgtgg	gaactgtaca	gtcatttctg	ctacaatgct	34440
gtatatgctc	ttctgaaaga	cacctatgca	aaatgcaca	gtaaaaatga	cacaactcat	34500
agggaaagcg	gggcccaggc	acagccctca	aaatctccat	caatgacatg	taagaaaaga	34560
gaggaacctg	ggaaatagca	aagtgccttt	tgcacattaa	atggttagct	atatcccaca	34620
atactgtgca	ttcgtaaacg	ttaatgctgc	aataaatacg	gcacttcacc	ttgggaagat	34680
ctggagttgg	cttatgagtg	tggaaagggg	tagcgcatga	gtttttgtga	aacactggaa	34740
ggaggattgt	gggaaatcaa	atggaaagtt	ctcaccacag	gcgtggagaa	gagtgggtca	34800
tggccccagc	agtgaagcca	gggagggtcag	agacggagggt	gtgtgtgtgg	gtgtgacct	34860
gcgcagttcc	ctgcgggtg	tagttttttg	cattcgctta	atgtttctcg	tggaggaaat	34920
tgtgcatgag	caaagtgtga	accgtgctgt	gtcacaattg	tcctaataca	tcattgcatt	34980
ggaacagatt	ggcttttttt	tttttttttt	tttttttttt	tttttgagat	ggagtctcac	35040
tctgtcacca	gcctggagtg	cagtggcatg	atcttggctc	actgcaacct	ttgcctccta	35100
tgttcaagtg	atcttctctg	ctcagcctcc	tgagtaactg	ggattacagg	catgagccac	35160
cgcggccggc	cagatttgca	tttttgaaac	aactgctagg	ctgggcgcgg	tggctcacac	35220
ctgtaatccc	agcactgtgg	gaggccgagg	cagggtggatc	acctgagggtc	aggggttcga	35280
gaccagcctg	gccaacatgg	tgaaaccccg	tctctactga	atatacaaaa	atcagctggg	35340
tgtggtaggg	ggtgcctgta	atcccagcta	ctcaggaggc	tgaggcagga	gaattgcttg	35400
aaccaggag	gcagaggttg	cggtagagccg	agatcacacc	attgcactcc	agcctgggca	35460

acaagagcaa	aactccatct	caaaaaataa	aaaatagaaa	aacaagtgtc	gtagcggaag	35520
tgagcacttt	gcgaggtcag	gcttgtgtgg	cctgttccac	aaatgatgtg	ctcacgggtg	35580
cctcaggccc	acctggagtc	tgagcatggg	ggcacaacag	gttcattagt	gtagaattcc	35640
aggacaggcc	tggctcctaa	gcagccttct	tttacaaaaa	ctgcagagcc	cgctgtatc	35700
ctagcacttt	gggaggccga	agtgggtgga	tcacgaggtc	aggagttaa	gaccagcctg	35760
gccaacatgg	tgaaccccca	tctctactaa	atatacga	attagctggg	tgtggtggca	35820
cgcgctgtg	gtcccagcta	ctcgggaggg	tgaggcagaa	ttgcttgaac	ctgggaggtg	35880
gaggttgtag	ggatctgaga	ccatgtcatt	gcactccagc	ctgggcaaca	gagcgagacg	35940
ccatctcaaa	aaaaaaaaaac	ctacagagcc	acacggcctc	tttctccacc	gagtgttggt	36000
gtgggagcct	gtgttattgt	ggtgaaatct	tggtactttc	ttgaggcaga	gagaggctga	36060
gcgcctggag	agactttcac	atgggtcgcc	atgtccgcgc	tcggtttcgc	tgttgtgtct	36120
cccactctgaa	ggctgggtgc	gtccagacag	gctggacgcc	cctttccacc	agatccttcc	36180
tcccgcagca	gtttctagtt	acgttgtact	gtgaggtctg	tgtccttggt	tgatggcaaa	36240
agtcagccga	attgaaattc	agagccatgc	ctggctccct	ggagcttctc	tcttgggcag	36300
ctgtgatcat	tgctctgtct	gtggtgtggg	tgggtgaaat	ggattccttt	catcttgctt	36360
gctacaggtg	actgtcacgt	ggagtccttt	ggagagaggg	acgtgttaat	tgatggatgt	36420
ggctcccag	ctgagaaagc	tcctgggcgt	acattgcctt	agagtttcat	tggagctgcg	36480
ttcttttatg	gtgtctgcta	ggcagaagtg	atgaagactt	ggaagaaaac	ccagaagggt	36540
ttccactaa	tttgaaaaat	gtgcttttcc	cctcctgtgt	cttttgctaa	ggtccagcct	36600
cctgcagcct	ccccgctctg	tggactctgg	ctttgattct	ttattaggag	tccccctgct	36660
cccccaaaag	atggtgtcta	aattatcatc	caattggccg	aggttttgtt	ttctattaat	36720
tgtttttatt	ttttattgtg	gtaaattttat	ataacataaa	atttgccatt	ttaattgttt	36780
tgttattgtt	gtttttgaga	cagggtctca	ccccagtgcc	caggctggag	tgcagtgggtg	36840
cgatcatggc	tactgagcgc	ctcagcctcc	agggctccag	tgatcctctc	acctcagcct	36900
ctctagtagc	cgggactaca	ggcatacact	accacatctg	gctgattttt	tgtatttttt	36960
ttttattgta	gagaccgct	atgttgccca	ggctggtctc	aactcctgga	ctcaagccat	37020
cctcccacct	cacctcccca	aagtgtctgg	attacaggca	tgagccacaa	caccagcca	37080
ttttaatttt	tttttttttt	tttgagatgg	agtctcactc	tatcgcccag	gctggagtg	37140
agtggcgtgg	tatcaactca	ctgcaacctc	tgcctcccag	gttcaagcga	ctctcctgcc	37200
tcagcctcct	cccagtagtc	tgggattaca	ggtgcccctc	actatgcctg	gctaattttt	37260
gtatttttta	gcagagacgg	ggtttcacca	tgttgccag	gctggtcttg	aactcctaac	37320
ctgggtgatcc	gcccgcctcg	gcctcccaaa	atgctgagat	tacaggtgtg	agccaccgtg	37380
ccgggccttt	ttttgttttt	gagacaggg	cttgccctgt	caccagact	ggagtgcatt	37440
ggtgggctct	tggctcactg	cagcctccgc	ctcccaggct	caagtgtgtc	acctccacac	37500
ctggctaact	gtattttatg	tagagacaga	tttcaccatg	ttgcccaggc	tgggcttgaa	37560
atggactcaa	gcagtcacc	cacctcagcc	tcccaaagtg	ctgagattac	aggcgcgagc	37620
caccgcaccc	agccattttt	acctattctg	cagttgacag	ttcagtgga	ttcagtcagt	37680
tcacgaggta	accatcactg	ccattcatct	ccagactact	tcaccttctc	ggcagatgtc	37740
cgaaactgtc	cgcattgaac	acactcctca	tctcctctg	acagccacca	ttctactttg	37800
tatctctctc	tgccttctct	aggtacctga	tgttaagtga	attataccaa	tatttgccct	37860
tgtgtgactg	gcttctttca	tgtgacatgg	gttcatctgt	gttatagcct	gttatagcct	37920
gtgtcagaat	ttccttctct	aaagcctgaa	taataaccgc	ttgtaaaggc	tgggcgcggt	37980
ggctcacacc	ctctaattcc	agcattttgg	gagtcaggag	tgggcagatc	acttgagggtc	38040
aggagtttga	gaccagcctg	gccaacatag	tgaaccctg	gctctactaa	aagtacaaaa	38100
ttagctgggt	gtgggtggcg	gcacctgtaa	tcccagttac	tcaggaggct	gaggcaggag	38160
aatcgcttgt	acccgggagg	cagagggtgc	agtgaaccaa	gattgtgcct	ctgcagtcca	38220
gcctgggtaa	cagagtgaga	cttcctgtct	caaaaaaaaa	aaaaatcatc	ggatggatgg	38280
acggaccact	tcttgttatt	tatccatcca	cgggtgctag	gtttcttcca	cctttggttg	38340
tcgtgaataa	ggccactatg	aacatttctt	tccgtggtga	aggttttgta	ctagttagga	38400
aaaggcgtgt	ttgtggtgtt	gcataggatt	ctggtgaagaa	agtgtgcact	aaccataagt	38460
atttgtacta	cattaaaatg	aaagctcagg	ggcggggcgc	ggtggctcac	gcctgtaatc	38520
ccagcacttt	gggaggccag	ggcggggcga	tcatgaggtc	aggagatcaa	gaccatcctg	38580
gccaacatgg	tgaaccccg	tctctactaa	aaataccaaa	aaactagcca	ggtgtggtgg	38640
cgggcacctg	tagtcccagc	tacttgggag	gctgaggcag	gagaatggcg	tgaacccggg	38700

aggcggagct	tgcggtgagc	cgagatcgct	tcaactgcact	cgagcctggg	caacagagca	38760
agactccgtc	tcacgcaaaa	ctctgtctca	cgcaagactc	cgtctcaaaa	aaaaaaagag	38820
ttcagggttt	atgaaactgg	ccagccgcgt	aaagtttgct	gtgttggttt	tgtgcccggg	38880
aggagtgtgg	ccagggtgtc	acgtcacaca	gtacacgttt	ctcagatggg	ggttctccag	38940
actgctgtcc	caaagtctgt	ttttgcatct	ggttcccaca	gacccaccct	ccacgggtgag	39000
cctgattttg	gccagggtag	ctggaatctt	gcttgctctt	cagcccggca	gctgtaccag	39060
tccagggtcc	acagctagt	gcttttagga	aggaatttgt	tcagttggct	ttgacacatg	39120
gccccctagg	gtccacagct	ctgtagtgat	gtggatgttg	ttatctacaa	agacacatga	39180
tccttcgtgt	ccagatgaaa	gtgatgatgt	ctttgcagct	gcccagcaag	gctgtgtgtg	39240
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgggtgtgtg	gtggtgtgtg	tgtgtgtatg	39300
ggggaggagg	gcaccctttc	catctggggg	tgtgtgtgtg	tggggtgtgt	gtgtgtgtgt	39360
gcgcgtgtgt	gtggtgtgtg	gtgtgtgtgt	gtgtatgggg	gaggcaccct	ttccatctgg	39420
gtccaagaga	ctgggcctgg	ggaagacgct	tctttttatc	tacttagaga	ctttgtttta	39480
tttgtatttt	tttgagacag	ggtctcactc	tgtcaccagg	gctgggggat	ggtgatatga	39540
gcatagctca	ctgcagcctc	ggcctcccag	gctgaagcga	tcctcccacc	tcagccttct	39600
gaatagctgg	gactgtaggc	gtgcgtcacc	atactgagct	attgtttttt	ttgtttgggt	39660
ggtttaattt	tttttgatac	agatggagtc	ttgctatggt	gcccagacta	gtctcaaact	39720
cctgaactca	agtgattctc	ccacctcagt	ttcccagcat	tctgggatca	caggtgtgag	39780
ccactgctgt	ctccctgttt	tattaactgc	tgaaagacct	agataaagaa	agtctgaaaa	39840
gacttactat	cagagcacca	tcctaagatg	attccctctg	actcaatgga	gagggagggg	39900
agcttttctc	tcaggcctgg	gtggcaggag	cccagggtgt	ccaggcccca	tttgccccag	39960
gccaaattcac	tcgggaactt	ggatgcagct	gtctttcagg	gtaaccctaa	ggaaccagat	40020
ccccgcaggc	agttaggttc	tgggctgtcc	tctcctccta	cgtcagctca	gtaagagccc	40080
ttcgaaggga	tgtgtgtgtg	gaggccccaa	aagcccaggc	tcacccctga	gatgcacagg	40140
gtgggctggg	cttaggcagc	gctcgagcat	ctcctggacg	gtgaccccag	agagtgtgga	40200
gacggagagt	ccttgagagt	cactgagaga	cgtggctgcc	ctgccttccc	aagaggggct	40260
ctgagtcatt	ccccacactc	acctgcccc	acccaccctc	acctggcccc	cagcctcacc	40320
tacccccaca	tctgtaccga	tccttttacc	cgcaccttcc	ctacccaccc	tcacctcccc	40380
tgtaccttca	cctccccac	tcacccgccc	ctgcaccctc	acctgtcccc	caccttcacc	40440
taacccccac	cctcacctgc	cctccccctc	cctggcctcc	ttccgttggg	gaaggggttg	40500
taaggggagg	cccccaaact	gtctgtcctg	gtgccctgca	gagaaaacag	tacgtgaggg	40560
ccgcagtcga	aaagcttgag	tcctggaagg	tggaggagac	agggatgtgt	tgggaagggc	40620
cccatgggtc	tggatccctt	ctcgactgtc	aatggggcct	tcattgggagc	gccagtctag	40680
tgatgcacag	ctgggtgccc	ggcggtggc	tgaggaggcc	taaagtccga	ggcggaaga	40740
gctcttccag	aggctgttgt	cctaactcgt	ctggcatact	caggcgggca	cgtagttagg	40800
agctgattgg	agaggagaga	ccccacacc	aatactggga	tttgactttc	aggctaaact	40860
tgagaagtgt	ggcctctgct	gtcctgccag	agctctccag	ccagtgccca	gggctctcca	40920
gccagtggcc	gggggtctcc	accagtggcc	gggggtctcc	gccagtggcc	gggggtctcc	40980
ccagtggcca	gggggtctcc	ccagtgtcca	ggagtcttgg	tttctttgtc	ttacagccct	41040
ttgttttgac	ctctctgagc	caaggccaaa	acccagacag	gcagccccc	gacctcagca	41100
tcgacatcta	cagccggaca	ctgttctgga	cgtgcgaggc	caccaatacc	atcaacgtcc	41160
acaggctgag	cggggaagcc	atgggggtgg	tgtctgctgg	ggaccgagac	aagcccaggg	41220
ccatcgctcg	caacgcggag	cgagggtagg	aggccaacgg	gtgggtgggg	gtgctgccc	41280
tcaggcgctg	ccgcgcgtgt	cttatgccga	atgccagcct	ctcacaggct	ggggagactt	41340
tcacactggg	gatccaatgg	gtggctttcc	agggtcccaa	aagcaaacac	aggtttttca	41400
cagcccgtcc	gggaaagcag	aaagccccaa	ggggctggaa	ggggaaaggg	ggagctctgc	41460
tgagagggtta	caaggcagcg	ctggccgacg	ggagttgcag	ttgatagggt	ttgtatcatc	41520
cttgttaaac	ttgaacctg	tgcagaaatc	ccttccacgg	catgggggct	gcctgttgac	41580
tgcctcctgt	tccaccacag	ggagctcctg	ggcttcttcc	tcccagaggg	ccccgacgct	41640
cccacctgtt	ggtcgtcaga	gcttctgggt	ggtgggaagg	cacccaggac	cttgaggtct	41700
ccagagagaa	aagccaggga	aagagggaga	ccgaaaccca	tgtgacatga	aactcaggct	41760
ccaaactgag	cacgggaacg	tttggggaca	ggagcgcgat	ggccttccct	agatagctgg	41820
ggggctggca	tgaagacggg	agctacagcc	agcacaggct	ctgggcccgg	agcccagaga	41880
ttgagccctg	actctgtcac	ttactggcca	cgtgaccttg	ggcggtggc	atagcctctt	41940

ggagactcag	tttcctcatt	ggtaggagt	acggccacag	tggcgcgcc	tctgcagcac	42000
acggggggct	cgggtggcgg	aagccccggg	tctataaggc	ggctgtgcag	gagccagccg	42060
agctgggtctc	ccaacagcca	gggctccggg	gtccttagca	gctgtggggg	gcctgcacct	42120
gtttcccatg	gctgctgtca	gaaattacca	gaagccaggt	ggctgagagt	aatggacact	42180
tgttctctca	cagttcctga	gggctgaagc	ccgagatcga	gggtgtgggca	gggccctgcg	42240
ccctctgaag	gctctgaggg	aacctttggg	cttctgggtg	ctccaggcac	cccttgactt	42300
gtggctcctgt	cactccagtc	tctctgtctg	gctgcacatg	gcgtggcctc	ttctgtacca	42360
ttgaaggaca	cttcagttgg	atthagggcc	tacctcacc	cattgtggtc	gtatcttgat	42420
ccttcctgac	atgtgtaaag	accctgcttc	caaataagct	cacattctga	ggttctgggg	42480
tgagcgggaa	tttgagagc	attgttcaac	tagtatagaa	tgtgacctgt	cagcctcggg	42540
cagccctgag	aggcaggggc	tttccacagc	ccagctgggt	gccctgggct	ccgtgctgtc	42600
cgaggagacg	ccatccccac	accgctcctt	caccgcccac	cctcccgag	gtacctgtac	42660
ttcaccaaca	tgcaggaccg	ggcagccaag	atcgaacgcg	cagccctgga	cggcaccgag	42720
cgcgaggtcc	tcttcaccac	cggcctcatc	cggcctgtgg	ccctgggtgg	agacaacaca	42780
ctgggcaagc	tggtctgggt	ggacgcggac	ctgaagcgca	ttgagagctg	tgacctgtca	42840
ggtacgcgcc	ccggggcctg	ccctaaccgc	agacaccggg	ccttcattgt	cagtaatggc	42900
agcagctgcc	acattgtccg	agacctgccg	tgagcccagt	gccgcgccag	gggctttgtg	42960
tgtagcgtgt	tttgcctca	cactgacagc	tgtaggctgg	ggttctgagt	gagccccaca	43020
gggcagaggg	agaaaatgag	tctcagagag	ggtgagcgag	ctgcttgggg	ccccacagca	43080
ggagatggag	caggactgca	gcctagcctc	tgccccagc	acctgcgcaa	gaagctgtc	43140
tgctctggac	tggtgtaggc	tgcgagggct	ggagagaaat	gagagttggg	gcttagagag	43200
ggggcgccag	tccccatggc	ttttcctctt	atgatgaggt	agatgggtga	agggaggggc	43260
catgcttgca	ggggccagtg	accgaggccc	gccgttggaa	ctgatggcct	tcatcccag	43320
cccagcccag	gtgggagcag	ggctttccga	gggcttgtct	tgggtcgccc	tgcttccagg	43380
gactctgctg	cagctcccac	ccctgtccaa	agcatggaat	ccccaggct	ccctggcagt	43440
cctgtcaacc	tctgtcctcc	caagctgagt	gtggggcaag	ttctggaggt	cagcactgct	43500
cagggggggc	cacgggctgc	ttgcaggggc	caaccgcctg	accctggagg	acgccaacat	43560
cgtgcagcct	ctgggcctga	ccatccttgg	caagcatctc	tactggatcg	accgccagca	43620
gcagatgatc	gagcgtgtgg	agaagaccac	cggggacaag	cggactcgca	tccagggccg	43680
tgtcgcccac	ctcactggca	tccatgcagt	ggaggaagtc	agcctggagg	agttctgtac	43740
gtgggggctg	gcagtggggg	gggcaggggt	gcctctaaac	ccgaccctcg	gaggaggctg	43800
gaggccagtg	caagatcctg	tgtggcctca	gccaggcggt	gggtctctgc	agatgccaac	43860
tggtgcccgc	tggggttcag	cgacatgtcc	gaatgtcccg	aggcctctga	ggttgttttc	43920
ttttgccgca	gaacaaatca	ccacgaacag	cgttttaaga	caacaccaac	tctttttttt	43980
tttttttttt	tgagtcagga	tcttgctctg	ttgcccaggc	tgggggtgcc	tggtgcaaac	44040
acagttcact	gcagcctcga	cctctgggct	taattaagtg	aacaccttgc	ctcagcctcc	44100
caggtagctg	ggactacagg	tgggcaccac	cacacctggc	taattttttt	ttgtagagac	44160
gggggtttccc	catgttgccc	aggctggctc	gcaactcctg	ggcacaagct	atctgcctgc	44220
tgtaggcctcc	caaagtgcta	ggattatagg	tgtgagccac	tggcctgaca	acaccacagg	44280
attgtctctc	agttctgtaa	ggcaaaagtcc	aggcacagcg	tggctcacct	gggttctctg	44340
ctcagggtct	cacggggcca	gaatcaaggt	gtcaggaacg	ctggggccctc	agcggaggct	44400
ctgtggagaa	attagcttcc	ttgctcactc	agcaggtagc	agttgtggga	tcgaggttct	44460
gttttctctc	tggttattgg	tcggggacca	ctctcagctc	ctagaggcca	ccacagggtcc	44520
ttgccccgtg	gccctctctg	cctcagcagt	gggggctccc	tgcgtcagtc	cctcccacac	44580
cttgagtctc	tctgatttgc	ttctaaaggg	ccctgtgatt	cggctcagcc	acctttagat	44640
taggttagcc	tcccctttga	tagactccaa	gtcggctgat	taataacctt	aatcacatct	44700
gcagaatccc	ttctgccaca	taaggctcatg	acgccgtgct	ggggactggg	gtgggaaatt	44760
acgggggtcat	ttaggattct	gcctgccact	gccttgctgt	gtcccagggc	ttggggggagg	44820
ggcctccaca	gctgggacca	cagtccttcc	tccctccat	ggtaaccatc	tgaggattac	44880
ttgagaccag	cctgggcaac	atggtgagaa	ccatcccta	caaaaaatac	aaacaaaaag	44940
ggaccaggct	gggcttgggtg	gctcatgcct	ataatcccag	cactttggga	gaccaagggtg	45000
ggctgatcac	ttgaggttgg	gagttcgaga	ccagcctgcc	caacatagtg	aaatcccgtc	45060
tctactaaaa	atacaaaaat	tagctgggtg	tggtaggcagg	cgcctgtatt	cccagctact	45120
ggggaggctg	aggtgggaga	attacttgaa	cctgggaggc	ggaagttgca	gtgagccaaa	45180

attacgccac	tgcactccag	cctagggcaat	agagtggagac	tccgtctcaa	aaaaaaaaaa	45240
ggggccagggg	tggtagtgtac	aaagagaccc	tatcccaaaa	aaaccgaaca	ctgaatcctt	45300
gagactgagt	aaggacactg	tgaaatTTTT	ctgggtgggg	cagggaaacag	agcgtcttct	45360
gtcattttctt	ccacctgggt	gtggtcagct	ctccctccaa	gctgcctcct	cttcttctca	45420
ttgtccgggt	gttgacacaca	tttggTTaaC	tggatagaat	aacgcgagtt	cccagggact	45480
tggTccattt	gctattttat	tttattttat	tttattttat	tttattttatt	tattttattta	45540
tttattttatt	tattgagatg	gagtttcggt	tttgtcgccc	aggctggagt	gcagtggcgc	45600
gatctcggtt	cactgcaacc	tctgcctccc	aggttcaagt	gattctccta	cctcagcctt	45660
ccaagtaact	gggattacag	gcacccacca	ccataccagg	ctaatttttt	tgtatttttta	45720
gtagagacgg	gttttcgcca	ttttgcccag	gctggtcttc	aactcctagc	ctcagggtgat	45780
ccacgcacct	cggcctccca	aagtgtctgg	attacaggca	tgagccacca	cgcctggcac	45840
catttgctat	tttaattccc	atgtgtattta	gtgtcccacg	gctgctgtaa	caaatagacca	45900
caaactggat	ggcttaaagc	aacagaaaatg	gattccccca	atgtgctgga	gaccagaagc	45960
ctgcgaccaa	actgttggga	gggctgtgct	tccctctggg	gctccaggga	ggatctattt	46020
gttgGCCctt	ccagtgtgt	gggtgccagc	gttccacact	tgtggatgcg	ccgcctcaac	46080
ctctgcccct	cttcatgtgt	ccatctcctt	tgtgtctgcg	tctttacctc	ttcttcttgt	46140
ctgtgttgcc	tcttataagg	acgtttgtca	ttgggttttag	ggcccaacca	aatcatccga	46200
gatgacctcg	tcttgagatc	cttaacctgc	aaagaccctt	tttccaaaaa	aaggttatgc	46260
tcacagattc	taggccttaa	gacatgggtg	tatctttctg	gggggcacta	tccaaccctt	46320
tatacaatga	aagacgggaa	gagggccagg	tgtggtagtt	cacgcctgta	atctcagcac	46380
tttaggaagc	tgaagcggga	ggatcacttg	agcccaggag	tttacaagta	gctaggcaac	46440
atgatgagac	cccatttcta	caaaaaagtga	aaaaaaaaaa	aaaaaaaaaa	aagccagggtg	46500
tggTggctca	cacctgtaat	cccagcactt	tgggaggctg	aggcaggcag	atcacgaggt	46560
caggagattg	agaccatcct	ggctaacacg	gtgaaacccc	gtctctacta	aaaatacaaa	46620
aaattatggc	cgggcgcagt	ggctcccgcc	tgtaatccca	gcactttggg	aggccgaggt	46680
gggtgaatta	caaggtcaag	agatcgagac	catcttggct	aacacggtga	aaccccatca	46740
agatcacaag	gtcaagagat	ggagaccatc	ctggctaaca	cggtgaaacc	ccgtctctac	46800
taaaaataca	aaaaattagc	cgggcattgt	agcgggcgcc	tgtagtccca	gctgctcggg	46860
aggctgaggg	aggagaatgg	cgtgaacccg	ggaggcggag	cttgccgtga	gccgagatcg	46920
ctccatgccca	ctgcaactca	gcctgggtga	cagagtgaga	ctccgtctca	aaaaaaaaaa	46980
aaaaaaaaaa	aaaaaaaaagaa	aattagccag	gcacagtggc	aggTgcctat	tgtcccagct	47040
acttggggagg	ctaaggcagg	agaatggcat	gaacccggga	ggtggagttt	gcagtgaagcc	47100
gagatcatgc	cactgcgctc	cagcctgggc	gatagagcaa	gactctgtct	caaaaaaaaaa	47160
agccaggcat	ggtggtgcat	gcctgtagtc	ccagctactc	aagaggctga	ggcaggagggg	47220
ttgttcgacc	cacggagatc	aaggctacag	tgagccatga	tcgcaccact	gccctccagc	47280
ctgggtgaca	gagtgtgacc	ctgtctcaaa	gtaagtaaat	aggaggagag	acaagtgggc	47340
agttcagact	gatggtatgg	gcacagtaga	gactggtgca	gacaggctgg	cctgtgatgt	47400
caagcaactt	ctgtaattgt	ttccggcatc	catttgtgtg	tcaatttccg	tgtcagtagg	47460
aagactctgt	aggctgccaa	gaggaataag	tgggaggatc	ctcccagaga	ggccgggcct	47520
gcaggagggc	cagttctcat	gagttctcat	tctggcccta	ccctccaggc	tgtggtctcg	47580
aggTgggaga	cagagcctga	cctctgtttg	tcttgttttg	tctttgcagc	agcccaccca	47640
tgtgcccgtg	acaatgggtg	ctgctcccac	atctgtattg	ccaagggtga	tgggacacca	47700
cggTgctcat	gcccagtcca	cctcgtgctc	ctgcagaacc	tgctgacctg	tggaggtagg	47760
tgtgacctag	gtgctccttt	ggggTgatgg	acaggTacct	gattctctgc	ctgctaggct	47820
gctgcctggc	atccttttaa	aatcacagtc	cctgtggcat	ccagtttcca	aagctgattg	47880
tgtcttcctt	tgcctcctt	tcttttctac	tatgtgcatt	cggTgctatg	aattttcctc	47940
taagtactgc	gtttcctgca	tctcacaaat	tttgttacat	tttcattttc	aggtagtttg	48000
aatattttta	cacttctcct	gagatgacat	ctttggctca	tgtgttattt	agaagtgttg	48060
cttagtttct	aaagagttgg	ggcttttcca	gctgtctctc	tgcactgat	ttctaattta	48120
attctactgt	agtctgagag	cttattttat	atgatttctg	ttattttaaa	tgtgttgggt	48180
gtggTgtttt	tgtgttatt	gtttttgtgt	ctttttgttt	tgttttgctt	cgtttgtttt	48240
gtttttgaga	cagtgtcttg	ctctgtcact	caggctggag	tgcaatggcg	cgatctcagc	48300
tcaccgcaac	ctctgcctcc	cgggttcaag	tgatcctctt	gcctcagcct	cctgagtagc	48360
tgggattaca	ggtgcacgcc	accataccca	gctaattttt	gtatttttag	tagagacggg	48420

gtttcaccat	ggtggtcagg	ctgggtctcga	actcctgacc	tcgtgatccg	cccacctcgg	48480
cctcccaaag	tgctgggatt	ataggcggtga	gccactgtgc	ctggccatta	ggtgtgtttt	48540
atcaccacagc	atcatgcagt	ttatcttgggt	gaatgtttctg	tgtactcttg	aaaagaatgt	48600
ggattctgct	ggtgttgggt	ggagtgttcc	agaaacatca	attagatcca	gttgggttaat	48660
agtgtctatc	agggtgtctc	tatccttcct	tcctgactgc	ctgcttgagc	tgtcagttat	48720
tgacaggggt	gtggagtctc	caactctaata	ggtggatttg	tttattttctc	ctagtagttc	48780
tatctttttc	tctccttcta	cccttgatcc	tcttctcccc	ctagggcttc	ctgggtgttag	48840
tggtgggaga	gtgggtagt	gaagaacctg	gacttttaggg	ccaaagaggc	caggggtcaa	48900
atcctggctc	tgtcacttcc	cagttgagtg	accctggctg	gtgcctgaat	ctctgtgagc	48960
ctccacttcc	tcctctgtga	aattgagagc	acttacctgg	caggctgtca	tgggcatcaa	49020
gtaacagggc	actccacctg	gaccctgaca	cgtgatgcac	aggaatgcca	gctgctatgc	49080
catgggtgtg	gcagtagtaa	taaagtgacc	atctgtatcc	tcaccacagt	gaagcctgtc	49140
cagggctttc	tctcctatgc	ccccatgcct	ccaggtggcc	ttggatcctg	ttgggtctgt	49200
gctctgctca	gcgacctttc	tcccgtggga	gttccctggg	gttcagcttc	atcctacaga	49260
cagcagcaca	cactggctgt	gcaccctttt	tttttttttt	tttttttttt	tgagatggag	49320
tctcgctttt	ttcgcgagg	ctgaagtga	gtgggtgtgat	cttggctcac	tgcaacctct	49380
acctcctggg	ttcaagtgat	tttccctgct	cacctccca	agtagctggg	attacaggct	49440
cccaccacca	cgcccggtca	atthttgtat	tttcagtaga	gatgggtgtt	caccatgttg	49500
gccaggatgg	tcttgaactc	ctgacctcag	gtgatccgcc	cacctcagcc	tcccaaagtg	49560
cagggattac	aggcgtgagc	caccacaccc	ggagtgcggg	ttgttttttag	cagtttgtct	49620
tgttcctgga	gagactgggt	cctgcccagg	agctcgggga	gtagggccgc	ggggtgctgc	49680
ctcacacctc	gagtttggcc	gtaagcagag	gggacatttt	gtgactgtcc	ccctcctgag	49740
cttcccagca	gcttttctcc	aagttacag	ccaaaagctc	aggtggattt	gcaacccaac	49800
ggtgtctgtg	cacctcccac	tgatgccga	actgcccctg	ccaagaaacg	gggcccgtcag	49860
aacgctgcac	taactgcagc	cttgggcctc	catgcctagg	gccatgccct	tccatccacc	49920
acccccctggc	ctgggcccctg	ggccctcctg	gctcgggaac	tccaggcccc	ttcctcacgg	49980
ctcgagagac	gtgtattttac	cgcacagggt	cttgtcattc	tcttgtggcc	tcttctccag	50040
ggagatcaca	gaaggacagg	gcctcactga	ggtctcggac	atggaccctt	tgatagtggc	50100
aggagccagg	ctgggcaaga	ggcgccacca	gtcacctcag	cagtgccatc	accaccgcca	50160
ttcagccctt	ccctgagccg	ggcgcgcccc	tggtctctggc	cccagtgtcc	cagttacagc	50220
tcacaggagc	ttgtggtgcc	cagcggctgc	ttctgattga	gagtcgaggt	cggaggcttt	50280
gggaggctga	gaggctgctc	ggtttcacaa	ctgctgaggg	agacttgggc	tccatctcag	50340
gtatgcccc	tgtcgccctc	aacctccagc	caccggctct	ccgtgtcccc	catggccagg	50400
cacggcttgc	agacatctgt	cgttggctcc	tctcagccgt	cgtgggctga	ccctggcacg	50460
tcctcctgtg	gctgagccca	gtggggacag	ctgcttcctt	ttattaccct	agaactctcg	50520
tctttgatca	ggccccctcc	cctatgccac	acagtccctg	tcactcgggt	gagcccagta	50580
gtcatgggga	aggcctgctg	gttccaaaca	tccaaaggct	tgcgtgcagc	atgacagctt	50640
gaaaccgatg	ttttttacct	tgatcagatt	tcagcttggc	gggggctttg	ctcagctttc	50700
agtgaggcct	gggccgattt	cccagcatcc	cctcctgagg	ccagcctctg	tttccctgtga	50760
ttttctgcac	aaagtgggag	ggaggagtcc	taggaaatgg	ggggccacct	cgaagcctag	50820
gcctcctctg	gcttctctgt	gccagtgcgc	ccacgctttg	tgtctgtgtc	cccagcccat	50880
gggactctgc	tattccctga	gtgctgcgc	atgccagacc	cgcactgagg	acgtggagcc	50940
ccgaggggca	ggatggcctc	catggtcaca	cgtaggaagt	ggcctccacc	ctccgatgat	51000
cctctccctc	ctccctttca	gcgcctcccc	cgggggtgtc	ctcagccctc	ctgcctgtgc	51060
tttgtcccgt	cttctgcagg	cgcctgggac	gtgctgacag	gtcctctgcc	ggctcctgcc	51120
ttgctatgcg	cacgctgggt	accacagagg	cctggccctt	cttctgtagc	agtcccacac	51180
ccgcaacagg	tgtggctgct	gaccacctgc	tttctgcccc	tctggctcctg	aggaggggcg	51240
agtgggcact	caggcgtggc	tgagcagatg	tgtgttgccg	ggaggaggaa	ggactgctcc	51300
agtcagggct	gaatttccca	cccggagcat	ttctgctgta	tttgggtgtg	cgcctgctgc	51360
ttaaagctct	gattcccagt	tggcacccct	tcccttctgc	attgaaaaac	atacggatgc	51420
atgtcttctt	gcagtgaatg	tgtattctcc	cagcctctct	tctgggttgg	ggctggagggt	51480
ggagcggcac	acaggagccg	cagcagtgga	ggatgtgcgg	gtgcagcacc	ccgtacagca	51540
gggatgccaa	acccgcgctg	agtcctctc	aacttctgct	ttgaagccca	gtcacgccat	51600
tgcttgggtt	ttgctggggc	gggctgcgtg	tgatgttctc	ctctgtccct	ccccagagc	51660

cgccacactg	ctccccggac	cagtttgcac	gtgccacagg	ggagatcgac	tgtatccccg	51720
gggcctggcg	ctgtgacggc	tttcccagat	gcgatgacca	gagcgacgag	gagggctgcc	51780
ccgtgtgctc	cgccgcccag	ttcccctgcg	cgcggggtca	gtgtgtggac	ctgcgcctgc	51840
gctgcgacgg	cgaggcagac	tgtcaggacc	gctcagacga	ggcggactgt	gacggtgagg	51900
ccctccccgt	caaggctctg	ccaagaccct	ggccctgccc	tccgggatac	gagcttgggg	51960
ctgcctccgg	cctcacagga	gtaggggctc	tgaaaacctt	tgcttgacgg	gagattgcca	52020
agtctgtctt	ttaggcccaa	caaggaaaac	tctgcagttc	cacccatcct	gtcccaccag	52080
gtagtgtggc	ttgaaggcag	actgtgaggg	tctatctcac	cttcctgcat	taggtcagga	52140
gtttcacaga	aacctgaggc	acattcaggg	gtgggctgca	gaggtccatg	gctcacaccc	52200
tggaaaatcc	gcccccaaaa	gacagtgtctg	tctccactga	ccagtctgtg	ggatagtgtc	52260
taagcctgag	tggtttctat	caacatgtag	aatcaggagg	tataaagaga	tttgctcagg	52320
catcctgggc	cctctctgac	cagcaggatc	ttccttttaga	tcttgacagt	gaaacacatc	52380
tcttctgtgc	cccctgtgag	ttttctttca	ttcattcatt	cattcattca	ttcattcatt	52440
cattcattcg	agacagagtc	ttgctctgtc	accaggcttg	gagtgccttg	gtgtaatctc	52500
ggctcactgc	aacctctgcc	tccagggttc	aatcgattct	cctgcctcag	cctcccaggt	52560
agctgggatg	acaggtgcgc	accaccatgc	ctggctaatt	tttgatattt	tagtagagac	52620
agggtttcac	catgttggcc	aggctggctc	cgaactcctg	acctcaggtg	atccgcccgc	52680
ctcagcctcc	caaagtgtctg	ggattacagg	catgagccac	cgcgcccggc	ctgagttttc	52740
cttttatgaa	ggacctgctt	ggttgggtgc	ctgccacatg	ttgtcagcac	catgggcccc	52800
ggactgtcga	ggagctgttg	atgccctcgc	tctcccagag	ccaccggctc	tgtagataaa	52860
ttcacatgca	gtctggccac	tgtcctacgt	cctcattcac	aaagagcaga	catttcgtag	52920
aagatgaggg	cctggggagta	acctccctgc	atgtttttct	ataaaggcat	agtgggttaag	52980
tccttccagc	tcattgacca	ttggagaatt	ttatggaggc	tgtagactag	gggctggtaa	53040
actaagggcc	caggggcccc	atccagcctg	ccacctactt	ttgtaaataa	agttttcttg	53100
gtgcacagcc	atgcccattc	attcatttgc	acaatgtctg	tggtgtcttt	catgccccaa	53160
gcaagagaac	tgagtggtta	tgtctggagac	ctacggcctt	caaagcccc	gacctcacgt	53220
ctggcccttg	acagacagag	cttccccagc	cctgctgcgc	atcctggccc	agcatgtgct	53280
gtgtgtgtga	tttcagcttg	caggagccgt	ggtaggaat	tgtccctgtg	ttgggtccatt	53340
ttgcattgct	atgaaggagc	acctgaggcc	gggtagatta	tgaaggaaaag	aggtctgtct	53400
ggctcatggg	tctgtaggca	gcaccagtat	ggcaccgcga	tctgctcagc	ttctagttag	53460
gtctcaggaa	gctttgactc	atggtgaaaag	togaagcggg	agcaggtgca	tcacatgggtg	53520
agagagggag	caacggagag	agagagagag	cgcctctccc	tcttgccctc	accttgagag	53580
gagatgccag	gctcctttta	gtaaccagct	cccatgtgaa	ctcacagtga	gagcccatth	53640
gctactgccc	agagggcacc	aggcatctgc	tcccatgacc	caaacactgc	ccaccaggcc	53700
ctacctccaa	ccttgggggtc	atatttttatt	ctgttctatg	ctatgctatg	ctatgccatg	53760
ccatgccatg	ccatgctatt	cctattctat	tatttgagac	agaatctcgc	tctgttgccc	53820
aggctggagt	gcagtggcat	gatcttggct	cactgcaacc	tccacctccc	aggttcaagc	53880
gattctcctg	tatcagcctc	cogagtagct	gggattacag	gcacacacca	ccacaccggg	53940
ctaatttttg	tattttcaat	agagatgggg	tttcaccatg	ttggccaggc	tggtctcaaa	54000
ctcctggcct	caagtgatcc	acctacctcg	gcctcccaaa	tgcccatgat	tacagatgtg	54060
agtcactgcg	cccagttagg	gtcacatttc	cgttgagatt	tggaggggca	gacgttggag	54120
ccatctgagc	cccctcgtcc	cgtcttagct	tctcctcccg	tgtgccccgc	gggtgctggg	54180
gcaggccctt	acgcgggttc	tggctgcacg	ctctgttcca	gaagctttct	tcctgcttg	54240
gttaccagaa	aatcatccca	tccattacaa	ggacagggtc	cccttatctc	ccattcccag	54300
ggcaggacac	cgggggcagg	gcagggtggg	aactgagcaa	gttctctggg	ggcaggcgtg	54360
gctatggctc	cctctgggtg	ggcgtctggg	gaggggtgga	ggcagccgtc	agcgccttg	54420
cttgcctctc	ctccctggcc	agagactgtg	gccttgtgct	gctcccgtgt	gggctgcctg	54480
cacctccagt	gggttgtgct	ccctccctc	ccctccctc	aagctctgct	gagcaccact	54540
gccttcacac	gccccactc	tggggaggcg	aggctcctcg	tggccattcc	tgtccttggc	54600
acccaccccc	ccaccaacct	ggttagagcct	tggggcgggt	ctgttactcc	ttgcatggcg	54660
tagacctccc	cacagttagg	acctgacaca	tacctcctgg	ggggcaggca	ggaggtgcgt	54720
tgaggtctca	gccctggcag	tccctccctc	gcgtggcata	ggcctcgcca	cagggtcac	54780
gaggggtggg	ggagactgta	ctagaccact	ccccgctggt	cctagaaaag	gtcccactctg	54840
tctgctctct	gtttggagtc	cagaccttgg	ttgctgtgcc	ctgcatgggtg	ggctgggggg	54900

caccctccag	cctctctgag	tgcattggcct	ctccttgag	ccatctgcct	gcccaccag	54960
ttccggtgtg	cgagcgcca	gtgtgtcttc	atcaaacagc	agtgcgactc	cttccccgac	55020
tgtatcgacg	gctccgacga	gctcatgtgt	ggtagagccag	cttctggcac	ggggaagggg	55080
cgtccgggct	gggttcccc	aggaacgtgg	agtttagggg	aggagacgtg	cctttccagc	55140
ggggctgggg	gctgtgtggg	agactcaggg	ggctgggagg	ctccttgagg	gaggcagggg	55200
agcctttccc	agggcagcgg	ccaggaggac	agactgtgag	ctgtgggctc	ggcggctaca	55260
gagtctgcct	cagtgggagg	ggctgatggg	gtccagggtg	ctgcagcacg	caccacacca	55320
cgggaccttg	ctgagcagcg	tctgtcaggg	agcaagatta	cccaggggct	gcagtgggtc	55380
tgttccctgg	cagcttactg	tctggctgag	gaggagtgat	gttcacatat	gcacacatgt	55440
catgtgcaca	cacatgtaca	tgacaacatc	ccacatgctc	ctcaaatagc	atgacctgta	55500
cagtacagga	tatagggcct	aggggatagg	aggccaagac	agtcagggaa	gactttccag	55560
aggcagtggc	tcttgaaagg	ctgtctgatt	caggcaggaa	gggagctgag	ttcagatagg	55620
aagtagcaat	gagtcattgt	gtctggggac	atggccactc	cttcgctgca	gagggaacctg	55680
ggctgagagc	tctctcttta	tggctgcagt	cgggagagaa	gtctgttggg	gggagaaggg	55740
ggcttcctca	agggactccc	tgtgcccctt	ggcaccttcg	tgccagggtc	ggcttgaggc	55800
ctgaaggcag	tgggtggggg	caccaagggt	cgcctcctct	gctgggcaag	ttcccagctc	55860
gacgggctcg	tgccgtgggg	cccagctgtg	ggggcgctgt	tgatgcgcag	ccaggcctcg	55920
ccggccagagc	cgcacgcctt	ccattccgct	gacttcacgt	acgcccctag	gatcgctggg	55980
ccggccctgt	gggagagtga	atgtggcctt	tgccaaaagt	gagtctggag	cctggaaact	56040
tccctatggg	cagccttgat	agtggagtgg	cccaaggagc	ccaccagacc	gacctgccc	56100
ctcccgctgg	tgggtggggg	caccaggggc	tgccctggct	tgctcgctca	ccaacatcac	56160
ctgggctggc	cagggcgccg	tcacttctgc	caccaccgag	ggccctgggc	gaaggagtga	56220
ataccaggct	gccttggcag	ggatgtgttg	agggctgtgg	ggagtcggac	agcggcgggg	56280
gtcagaggag	gaggagggg	caccgtgcag	gctgaagggc	cacgttacc	tgagggtggc	56340
caggctcccc	aggcctagcc	tcccagctcc	cccactttct	ccccaccctc	caccagtggc	56400
aaagccagcc	ccttcagggc	gcacggtgtc	tgcccccaag	gagggcccat	tccgttgggg	56460
ttaatgttgg	ccacctcttt	ctgtttgtct	ctggcagaaa	tcaccaagcc	gccctcagac	56520
gacagcccgg	cccacagcag	tgccatcggg	cccgctcatt	gcacatcct	ctctctcttc	56580
gtcatgggtg	gtgtctatct	tgtgtgccag	cgcgtgggtg	gccagcgcta	tgccgggggg	56640
aacggggcct	tcccgcacga	gtatgtcagc	gggaccccg	acgtgcccct	caatttcata	56700
gccccggggg	gttcccagca	tggcccttc	acaggtaagg	agcctgagat	atggaatgat	56760
ctggaggagg	caggagagta	gtctgggcag	cttgggggag	tggagcaggg	atgtgctacc	56820
ccaggccctc	ttgcacatgt	ggcagacatt	gctaactgat	cacagcattc	agcctttccc	56880
actgagcctg	tgcttggcat	cagaatcctt	caacacagag	gcctgcatgg	ctgtagcaac	56940
ccaccctttg	gcactgtagg	tgtggagaaa	gctccttgga	cttgaccttc	atattctagt	57000
aggacatgtg	ctgtgttgtc	cacaaatcct	catgtaccct	agaaatgaat	gtggggggcg	57060
ctgggctctc	tccagagctg	aaggaatcac	tctgtaccat	acagcagctt	tgtcttgagt	57120
gcagctggga	tttgtggctg	agcagttaca	attcctacgt	ggcccaggca	ccagggaacgc	57180
aggctgtgtt	tgtagatggc	tgggcagccg	caccgcagag	ctgcaccatg	ctgggtttgt	57240
tcacatgggt	gacctgggt	tgtctaagaa	ggtaggtcc	ctgtgaggtc	tgccaggtgc	57300
cccacagctc	caggccacct	tgaggattgc	ctctgcctgc	ccagccctga	gttccctctc	57360
ccctgtcctg	tcccactgtc	accccaagcc	ggcctcattg	ggagcctgtt	ggatggcagg	57420
gtatagatgt	aacctgatct	tctctgggga	gcgggggtat	ctggcttctc	aagagctcct	57480
aggagcccac	agtgggtggc	ccatcacagt	cgcagcagcc	cccagagaa	gcggccctgt	57540
ctgttccctg	cgtgctctgt	gctgccccgc	ctgggttccc	tgccccagtc	gcaggccctc	57600
tggaggagg	accatgtgtc	tcccgtttca	cagatgagcc	ccggggagct	cactctagta	57660
gtggccagag	aggcctgcgg	ctcaggagc	ggggcacatt	tccaacagga	cacaccgccc	57720
tggctctgag	ctcgtgggta	gtgggagcag	aggagagcgc	cctatgtctg	tggggcggct	57780
tggctgagcc	tgggaagccac	ctgacctccc	cctgcccctc	cctgcccagg	atcgcatgcg	57840
gaaagtccat	gatgagctcc	gtgagcctga	tggggggccg	gggcgggggt	cccctctacg	57900
accggaacca	cgtcacaggg	gcctcgtcca	gcagctcgtc	cagcacgaag	gccacgctgt	57960
accgcgggt	gagggggcgg	gccggggagg	ggcggggcgg	gatggggctg	tggggccctc	58020
ccaccgtcag	tgctggccac	cggaggcttc	ccgggttcc	gggggctgtg	ccaccgcctc	58080
tgaggcatgc	ttgctttctt	cccttttcaa	acccttctgc	ttccttcttt	aatgacattg	58140

ttgattgtgg	ataatctgaa	aactacacaa	aatataaag	agccaaaatc	tcacccaaat	58200
ccacctccta	gagtggctgt	tgggctccgt	cagcatccag	gcggccgtct	gtgttccgca	58260
cgcccagcc	catogatagc	cgctgcacc	aggcctgtct	gccctctgtg	agcctcccca	58320
cagggttccc	tccacaaaca	cctgtttctc	cccccagggy	ctggctgctt	cctggaaaac	58380
agctggatgg	ttttgtgcat	gacagacaaa	cacagggtga	ttttcgtggc	taaaatactc	58440
cctggagctt	ttggcagggg	gaggggctgg	ctccagctga	gccacgcctt	gagtgaatg	58500
actgtgagga	gaataaactg	ccgctgccct	ccaggatcac	tggggctggc	tggggagaac	58560
ccccgtttct	gggagcacag	tcccaggatg	ccaaggcgag	cttgggtgcc	agatgtgaac	58620
tcctgagtgt	aaacagcggg	ggctgacttg	acatgctttg	tatgcttttc	atltgttcoct	58680
gcagctgtat	gcccctaagg	tgagtccagc	ccccctctgc	ttcctctggg	gcctcgccag	58740
tgagccccac	cttgcctggg	ctggttcctc	ctgcccttct	gggtatccct	cacatctggg	58800
gtcttgtctt	cttgttttct	ttttcttttt	tttttgagac	ggagtttcac	ttttgttgcc	58860
caggcttcag	tgcaatgggt	tgatctctag	gctcacgcga	acctctgcct	cccaggttca	58920
agcagttccc	ctgcctcagc	ctccctagta	gctgggatta	caggcatgtg	ccaccacgcc	58980
cagctaattt	tgtattttta	gtagagatgg	ggtttctcca	tgttggtcag	gctgatcttg	59040
aactccctac	ctcaggtgat	ccgcccacct	tggcctccca	aagtgcctgg	attacaggcg	59100
tgagccaccg	cacctggcct	ttttcttttc	ttttcttttc	tttttctga	gacagggtct	59160
cgctctgtca	cccaggctgg	agtgcattgg	tgtcatcatg	gctaactgca	gcctctacct	59220
tctaggctca	agcaatcctc	ccatctcagc	ccctaagtag	ctaggactgc	acgcctgcct	59280
ccccatgccc	agctaataatt	tacatttttt	gtagagatga	agtttacta	tattgccag	59340
gctggctctc	aactcctgga	ctcgagcgat	cctcctgcct	cgccctcccc	aggtgctggg	59400
attacaggcg	tgagccaccg	tgcctggcct	ggggtattgt	cttcttatgg	cacctgactg	59460
tggtgggccc	tgggaaggaa	gtagcagaag	agggttcttc	ttggtttctc	ggacagtaac	59520
tgagtgttct	ggaggcccca	gggcctggct	ttgtttaggg	acaaaggga	ctggtaacca	59580
gaagccgaga	gtttaaacac	ccactgccct	tcttccctgc	tcctgctgct	gcaacccagc	59640
ttaaccagcc	aggagtgcta	ggaacccaag	cagggccccc	gagcacacag	caggcagctc	59700
acgaattctc	ttttcctggt	ctcccttggg	agctgggagg	atcttaata	ggcaataaga	59760
gatggcactg	agcagccagc	taatttttta	aatcacttta	ttgtttaacc	atatgactca	59820
cccacttaaa	aaagggtaca	gttcagtggg	ttttagtgtg	ttcacagatg	tgtgcaaccc	59880
tcaccacagt	taatttttaga	acattttcct	gcccctaaaa	gaaactctgc	atgaagccag	59940
ctgtttttta	attagcaaag	ttatttttga	tccttttaaat	atatgttcat	ggtacaaaat	60000
tcaaaagata	cagaagagtc	tgcagtccaa	agagactccg	cccccatgac	gccaagcagg	60060
catccctggg	aggcatggcc	tcctgcagtg	tgtttcttct	atgtccccc	aggggtcatc	60120
tgtacatatg	caagcataca	agagcgtgga	ctttgttttc	caagccagaa	gataattgta	60180
gatttatgtg	cagttgtgag	aaagagcaca	gacccattta	tcctctgcct	ggtttccccc	60240
agtgtctgct	gccatcttgc	atgacttcca	ttcctatcat	aagcaagaca	ctgataacga	60300
ttctttccac	ttattccagat	tgacataagt	tttttttggg	tgttcttgag	acaaacttcc	60360
tctgtcacc	agtgggagtg	cagtggcaca	atcacagctc	actgcagcct	caaactcctg	60420
ggctcaagcg	attctcctgc	ctcagtcccc	tcaagtagct	cagatggcag	gtgtgcacca	60480
tcattgccagg	ctaatttttta	aattttttgg	ggagggtagg	cctcactaaa	tttctggggc	60540
tagtcttgaa	ctcctgagct	aaagtgatcc	tcctgcctca	gcctcccaa	gtggtaggat	60600
tacaggcatg	agccactgcg	cctgggctga	catatgtgtt	ttcgtaagcc	cgaaagatag	60660
catctgaaga	gtcaacattg	agccttgcc	tttgcctgta	acgatgtata	aaagctgctg	60720
ttctgagcat	ttcggaggct	cccagctgcc	gtgtgcaccc	tgccctagagc	tctaccgtaa	60780
cccatctccg	ggaggagggtg	ctattgtttt	cctcattttg	caacaaggag	gctgaagaac	60840
tgagcatgaa	ccactggcct	gggtcgttcg	gttggtaggc	agtggggcca	ggccatccaa	60900
ctcacaca	ccttctactc	tgttccccc	gcaccctgaa	gtttgttctg	ttttgaggac	60960
acagccgtca	cattcttggg	ggctgaacag	cactccttgt	caggcgtggc	tgggccccca	61020
ctggaggggca	tcatggctct	ctctcctgct	gcgggttgaa	cttggctgtt	tcaaccactc	61080
ctgccaagtg	gcccctctgaa	agggacagtc	catcttttct	cagcagaggg	ccacactggc	61140
aaaacgggtcc	ctggcaccct	ttctctccac	ctgtctaata	tagagtaaaa	atggatcat	61200
gttaagatct	tcatttatat	ttattttatc	atgaatgatg	taagcatcat	tttgtgtgtt	61260
taagaacctt	tgggcccagc	gtgatggctt	gcagctgtaa	tctcagcact	ttaggaggct	61320
gagatgagcg	gatcacttga	ggccgggagt	ttgagaccag	cctggccaac	atggagaaac	61380

cccgctctcta	gtaaaaaattt	aaaaattagc	cggggtatggt	gatcccagct	acttgggagt	61440
ctgaagcatg	agaattgctt	gaacatggga	ggcggagggtt	gcagtgagcc	gagatcgccg	61500
cattgcactc	cagcctgggc	gacagagcga	gactctgtct	caaaaaaaaa	aaaaaaaaaag	61560
aaaagaaaag	aaattatcaa	tctcctcttt	tatggcatat	atatatatat	atatatatat	61620
atatatatat	atatatatatt	tttttttttg	gttatgttca	gaaaggcctt	cctgctctg	61680
atcataaaaa	acaacttatt	ttcacactct	ctctcttttt	tttttgagac	agagttttgc	61740
tcctgttgcc	caggctggag	tgcagtggcg	caatctcage	tcactgtaac	ctccgcctcc	61800
cggtgttgag	tgattctcct	gccttacctt	cccagtagc	tgggattata	ggcatgcacc	61860
accatgcctg	gctaattttg	tacttttagt	agagacgggg	gtttctccat	gttggtcagg	61920
ctggtctcga	actcgcgacc	tcagggtgatc	cacccacctc	ggcctcccaa	agtgtcggga	61980
ttacagacgt	gagccaccat	gcccagccca	cactctcttt	cttaacgtcc	tcctcctttc	62040
gttttacgtt	cacatcttta	attcttctgg	gatgtaatta	gatttgatga	gcaagggtggg	62100
catccagctt	gtttcttggt	tgatggctta	tgggtggcgt	gaattagtcg	gggtctatca	62160
ggaggcagaa	actctatgag	aatttgaaca	gagaaagttc	cgtctacagg	cttattacca	62220
gggactggaa	tagcagaaat	tgaacagtga	gatgtacaga	gaactctaag	aatgcaggaa	62280
taggccaggc	atggtggctc	acacctgtca	tcccagcact	ttgggagacc	aaggcgggtg	62340
gatcacctga	ggtcaggagt	tcgagaccag	cctggccaac	atagtgaaac	cccatctcta	62400
ctaaaaatac	aaaaaaatta	gctgggtgtg	gtggcgcatg	cctgtaatcc	cagctactcg	62460
ggaggctgag	gcaggagaat	cacttgaacc	tgggaggcag	agggtgcagt	gagccgagat	62520
catgccactg	tactccagcc	tgggtggaag	agcggaaactc	tgtctgaaaa	aaaaaaaaaaa	62580
aacaagaagt	tcaacttgaa	gggaaaaatg	ccgtattgtc	tttccctttg	ttatgtcacc	62640
agggcacagt	ccatcccagg	ctggcgctga	tccacgggct	ggagagggggc	tgccccagaa	62700
gaggacatgc	caggaaagggc	ttggctgggtg	ttcaggagcc	caggccagggt	cagggtcaaga	62760
ggtgttgagg	ctggacggga	gaggccagct	aggggctcat	gtaggatatg	aggggtcggc	62820
ccatttcaac	gtggaaactg	agctcttctg	cttctctttc	ttcttctactg	cattaagatt	62880
caataccgct	tgggaagcag	gtatttccct	tcctataaaag	gatgggtggg	agcctgagtg	62940
ttgggagaaa	gtgtagccgc	tgagttacta	acaactaggg	ctgccgtcaa	gcctatgggg	63000
aaagagagaa	gaggacattt	ggaaggagag	agatcaagct	gtggcaccct	gggagaggac	63060
cacagaaaag	aggccagtga	gggggttccc	cgggtggcatc	tgaagggtgtg	gccaaccag	63120
gaggtccaga	ggctgccagc	cgagtggccc	aggagaggga	acctcacagg	ggctgagtg	63180
gacccaagcc	ctatccaccg	tcctaaccac	ccacatttct	cgggaacaag	acctcccaca	63240
gtggcctccc	cggcagtgga	aatagccaaa	ctggcaacat	ggactttctt	caactgccc	63300
ggcgatgctg	cctcagtgcc	ccagggcagg	caggaagctc	ccacacccat	tctggaatga	63360
ggggttggag	gaaggctgag	ctgagcaaag	gacccatctc	tgctctggtt	ggtggggagg	63420
gagcccatta	tacaagagac	ccctcagggc	tcagtgaagg	gtgacagaga	cttggggagt	63480
agtggctgtc	actgcagagg	tgagagggtt	tggagagaag	gtacatgcct	ttttggccac	63540
attgagtagc	acctggtagc	cagttagtaa	cgtgtattgg	ataaacaaaa	gattaaacgg	63600
atgcaaaaaa	aatgtttggc	tttgcttctt	tttaccctaa	cctcagttcc	ctcaagtaga	63660
ttctgggaac	acccctacc	tggtggact	gttgtgaagt	ttaaataagc	cagggttaact	63720
tcacctctc	ctttaagaca	cagctcagac	actgcctcct	ccaagaagcc	ccctctggct	63780
tcctgtgtga	atatgacggc	cctctgggct	ctagggtatc	ttagaacaat	gcttccttat	63840
ggcttttgaa	ccccgctgtc	tcctggattg	ggagcaaatg	caggggagga	gccacacctg	63900
actaatctct	gggtctccca	gcacataagt	ggcataagg	cagggctgtg	cccgttccag	63960
gcacttactg	aaggatgtac	ttggcagagg	gtaggcagcc	ggcggatgag	ccctcactc	64020
tcccagctg	actgcgtggg	cgggaaaggc	gggttcagga	gacccagcct	ccctgggctg	64080
tcaccacctc	tgacatcca	gccccattga	tcaagggttc	aatttttggg	gtcctgttgg	64140
gaggccagga	gactctctcc	aggcacttct	tccaggtctt	tgtgttaggg	tgtgtgtgtg	64200
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtt	gttttatttt	atttatttat	ttatttattt	64260
atttatttat	ttatttattt	tgagacgcag	tctcgctctg	ttgcccagg	tggaggggtg	64320
tggcatgatc	tcggctcact	gcaagctccg	cctcccgggt	tcacgccatt	ctcctgcctc	64380
actcttcctg	agtagccgga	ttacaggcgc	acgcaccatg	cctggctaata	tattttgttt	64440
tttttagtaga	gacagggttt	cgccacgttg	cccaggctgg	tcttgaatcc	ctggcctcaa	64500
gcgatccgcc	cgcctcagcc	tcccaaagtg	ctgggattac	aggcgtgagc	caccgtgccc	64560
gcccagccta	ggggtacatg	aaactttttt	tttttttttt	ttgagacaga	gtttcactct	64620

```

gtcctcaggc tggagtgcag tggcgtgac tggcgtact gcaatctccg cctcccgggt 64680
caagegattc tctgcctca gcctcccag tagctgggat tgcaggcacg cgccaccaca 64740
cccagctaat ttttgatatt ttagtagaga cgggctttca ccatgtggga caggatggtc 64800
tcgatctcct gacctcgtga tccgcccgc tcagcctccg aaagtgtctg gattacaggc 64860
ctgagccacc gtgcccagcc atgatgtttt gatacaggca tataacgtat aataatcaca 64920
tcagggtaaa tgatgtaacc atcacatcaa gcatttatcc tttgtgttac aaaaaaaaaa 64980
ctaattatac tttcctactt attctttttt tttttttttt ttgagacgga gtctccctca 65040
gtcgcgccagg ctggagtgcg gtggcatgat ctacgttcac tgcaagctct gcctcctagc 65100
tctgcctcct ggggttcattg cattctcctg tctcagcctc gcgagtagct gggactacag 65160
gcgctgcca ccgtgcccgg ctaatttttt tttttgtatt tttggtagag acagggtttc 65220
accgtgttag ccaggatggg ctgatctcc tgacctata atccgcccgt ctcgccctcc 65280
caaagtgtcg ggattacagg catgagccac cgcgccagc ctatttattc ttaaagtac 65340
aataaattat tgttgactcc agtcaccctg ctgtgtctac aaatacggat cttcttcatt 65400
ctatctaaat gtatttctgt acctgttaac catctctcct ccacctcacc ccccaaacc 65460
actacccttc tcagcctctg gtaaccatcc ttctactctc tatctctatg agttcaattg 65520
tattaatttt tagctccccg gccgggcacg gtggctcacg cctgtaatcc cagcacttca 65580
ggaggctgag gcagggtgat cagaggtgag ggagtttgag accagcctgg ccaacatgg 65640
ggaaccccat ctctactaaa aacacaaaaa ttagctgggc gtgggtgggt gcgctttag 65700
tccagctac ttggggaggc gaggcaggag aatcgcttga aactgggagg cagaggttgc 65760
agtgagocaa gattgcgcca ctgcactcca gtctgggtga cagagtaaga ttccatcccg 65820
aaaaaaaaa agtttagctc ccacaaataa gtgagaacac gtgaagtttc tctttctgtg 65880
cctcgcttgt ttactttaac ataatgacct ccagttccat ccacgttgtt gctttgttat 65940
aaatgacagg atcttggtea ggcgcagtgg ctcatgcctg taatcccagc actttgggag 66000
gctgaggttg actgatcatg aggtcaagag atcgagacca tcttggttaa cacagtga 66060
ccccgtctct actaaaaata caagaaatta gccgggcgtg gtgggtgggc cccatttccg 66120
ccccctctcg ggacgtgat gcacgacata ttacccatcc ccggaagact aatcctcccc 66180
cactctatat tgtacctctt cctttctcct ccacgcgatt ccccgagtaa cccgtcttcc 66240
ctccctcctc ggattacgtt cacccttccg cttcaatcac gttgctccgt ccccttcccc 66300
attcgtaacca ctctcactt tctgttctc accccacta tcccttttct tctctctat 66360
tccttaactta ctctccccc ttctcttcat acttcattcc ctccgctctt cccactcgcg 66420
ctcccaactt cactagttg cctcaccta cgttgccatc tcgccccttc ttcagctctc 66480
ggcctctcac ccatctgtcc tctctcttac ctctctctc atctcgctca gacatctctc 66540
tagactatcc ctcactttac cttctcagtc gtcttcttcc tctcttctgt tctccatgat 66600
cttcacgtcg ccatctcttt tcgccccttt catatgtctc tcttcatgtt ctactatca 66660
ttctcatgat cactatcgtt ctcaactatt atcactcccc tctttcttca tcaattctc 66720
tccgtcattc tctctctct cttacaaccg ccttcttct gctatctaac tcaaccatgc 66780
ctctctact ctctctctat cgcctctcca tcgcttatgc atcctctct attgcacacc 66840
cgccctcca tcgcttatgc atcctctct attgcacacc gccctccat cgcttatgca 66900
tcctcttcta ttgcacatcc tcttctattg cac 66933

```

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 12

ctgagcggaa ttcgtgagac c

21

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 13

ttggtctcac gtattccgct cga

23

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 14

ctcgagaatt ctggatcctc

20

<210> 15

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 15

ttgaggatcc agaattctcg ag

22

<210> 16

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 16

tgtatgcgaa ttcgctgcgc g

21

<210> 17

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 17

ttcgcgcagc gaattcgcac aca

23

<210> 18

<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 18
gtccactgaa ttctcagtga g 21

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 19
ttgtcactga gaattcagtg gac 23

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 20
gaatccgaat tcctggtcag c 21

<210> 21
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 21
ttgctgacca ggaattcgga ttc 23

<210> 22
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 22
cuacuacuac uactgagcgg aattcgtgag acc 33

<210> 23

<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 23
cuacuacuac uactcgagaa ttctggatcc tc 32

<210> 24
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 24
cuacuacuac uatgtatgcg aattcgctgc gcg 33

<210> 25
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 25
cuacuacuac uagtcactg aattctcagt gag 33

<210> 26
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 26
cuacuacuac uagaatccga attcctggtc agc 33

<210> 27
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 27
aactggaaga attcgcggcc gcaggaattt tttttttttt ttttt 45

<210> 28
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 28
aattcggcac gag

13

<210> 29
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 29
ctcgtgccg

9

<210> 30
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 30
gtacgacggc cagt

14

<210> 31
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 31
aacagctatg accatg

16

<210> 32
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 32
ccaagttctg agaagtcc

18

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 33
aatacctgaa accatacctg 20

<210> 34
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 34
agctgctcgt agctgtctct ccctggatca cgggtacatg tactggacag actgggt 57

<210> 35
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 35
tgagacgccc ggattgagcg ggcagggata gcttattccc tgtgccgcat tacggc 56

<210> 36
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 36
agctgctcgt agctgtctct ccctgga 27

<210> 37
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial sequence is a primer.

<400> 37
gccgtaatgc ggcacaggga ataagct 27

<210> 38
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial sequence is a primer.

<400> 38
 gagaggctat atccctgggc 20

<210> 39
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial sequence is a primer.

<400> 39
 acagcacgtg tttaaagggg 20

<210> 40
 <211> 163
 <212> DNA
 <213> Homo sapiens

<400> 40
 actaaagcgc cgccgccgcg ccatggagcc cgagtgcgct cggcgcgggc ccgtccggcc 60
 gccggacaac atggaggcag ctccgcccg gcccgcgtgg ccgctgctgc tgctgctgct 120
 gctgctgctg gcgctgtgcg gctgcccggc ccccgccgcg gcc 163

<210> 41
 <211> 419
 <212> DNA
 <213> Homo sapiens

<400> 41
 gcccacagc ctgcgcgctc ctgctatttg ccaaccgccc ggacgtacgg ctggtggacg 60
 ccggcggagt caagctggag tccaccatcg tggtcagcgg cctggaggat gcggccgcag 120
 tggacttcca gttttccaag ggagccgtgt actggacaga cgtgagcgag gaggccatca 180
 agcagaccta cctgaaccag acggggggccg ccgtgcagaa cgtgggtcatc tccggcctgg 240
 tctctcccga cggcctcgcc tgcgactggg tgggcaagaa gctgtactgg acggactcag 300
 agaccaaccg catcgagggtg gccaacctca atggcacatc ccggaagggtg ctcttctggc 360
 aggaccttga ccagccgagg gccatcgctt tggaccccg ctcacgggtaa accctgctg 419

<210> 42
 <211> 221
 <212> DNA
 <213> Homo sapiens

<400> 42
 ccccgtcaca ggtacatgta ctggacagac tggggtgaga cgccccggat tgagcgggca 60
 gggatggatg gcagcaccgc gaagatcatt gtggactcgg acatttactg gcccaatgga 120

ctgaccatcg	acctggagga	gcagaagctc	tactgggctg	acgccaaagct	cagcttcatc	180
caccgtgcc	acctggacgg	ctcgttcgg	taggtacca	c		221

<210> 43

<211> 221

<212> DNA

<213> Homo sapiens

<400> 43

tccctgactg	caggcagaag	gtggtggagg	gcagcctgac	gcaccccttc	gccctgacgc	60
tctccgggga	cactctgtac	tggacagact	ggcagaccgg	ctccatccat	gcctgcaaca	120
agcgactgg	ggggaagagg	aaggagatcc	tgagtgcct	atactaccc	atggacatcc	180
aggtgctgag	ccaggagcgg	cagccttttt	gtgagtgccg	g		221

<210> 44

<211> 156

<212> DNA

<213> Homo sapiens

<400> 44

tttctcagtc	cacactcgct	gtgaggagga	caatggcggc	tggteccacc	tgtgcctgct	60
gtccccaagc	gagccttttt	acacatgcgc	ctgccccacg	ggtgtgcaga	tgcaggacaa	120
cggcaggacg	tgtaaggcag	gtgaggcggg	gggacg			156

<210> 45

<211> 416

<212> DNA

<213> Homo sapiens

<400> 45

ctccacagga	gccgaggagg	tgctgctgct	ggccccggcg	acggacctac	ggaggatctc	60
gctggacacg	ccggacttca	ccgacatcgt	gctgcagggtg	gacgacatcc	ggcacgccat	120
tgccatcgac	tacgacccgc	tagagggcta	tgtctactgg	acagatgacg	aggtgcgggc	180
catccgcagg	gcgtacctgg	acgggtctgg	ggcgcagacg	ctggtcaaca	ccgagatcaa	240
cgaccccgat	ggcatcgccg	tgcactgggt	ggccccgaaac	ctctactgga	ccgacacggg	300
caaggaccgc	atcgagggtga	cgcgcctcaa	cggcacctcc	cgcaagatcc	tgggtgtcgga	360
ggacctggac	gagccccgag	ccatcgccact	gcacccccgtg	atggggtaag	acgggc	416

<210> 46

<211> 198

<212> DNA

<213> Homo sapiens

<400> 46

ttctttctcca	gcctcatgta	ctggacagac	tggggagaga	accctaaaat	cgagtgtgcc	60
aacttggaatg	ggcaggagcg	gcgtgtgctg	gtcaatgcct	ccctcgggtg	gcccacggc	120
ctggccctgg	acctgcagga	ggggaagctc	tactggggag	acgccaaagac	agacaagatc	180
gaggtgaggc	tcctgtgg					198

<210> 47

<211> 244

<212> DNA

<213> Homo sapiens

<400> 47

ccgtcctgca	ggtgatcaat	gttgatggga	cgaagaggcg	gaccctcctg	gaggacaagc	60
tcccgacat	tttcgggttc	acgctgctgg	gggacttcat	ctactggact	gactggcagc	120
gccgcagcat	cgagcgggtg	cacaaggcca	aggccagccg	ggacgtcatc	attgaccagc	180
tgcccgcact	gatggggctc	aaagctgtga	atgtggccaa	ggtcgtcggg	gagtcggggg	240
ggtc						244

<210> 48

<211> 313

<212> DNA

<213> Homo sapiens

<400> 48

gttcgcttcc	aggaaccaac	ccgtgtgcgg	acaggaacgg	gggggtgcagc	cacctgtgct	60
tctgcacacc	ccacgcaacc	cgggtgtggct	gccccatcgg	cctggagctg	ctgagtgcaca	120
tgaagacctg	catcgtgcct	gaggcctttt	tggtcttcac	cagcagagcc	gccatccaca	180
ggatctccct	cgagaccaat	aacaacgacg	tggccatccc	gtcacgggc	gtcaaggagg	240
cctcagccct	ggactttgat	gtgtccaaca	accacatcta	ctggacagac	gtcagcctga	300
aggtagcgtg	ggc					313

<210> 49

<211> 255

<212> DNA

<213> Homo sapiens

<400> 49

cctgctgcc	gaccatcagc	cgcgcttca	tgaacgggag	ctcgggtggag	cacgtgggtg	60
agtttggcct	tgactacccc	gagggcatgg	ccgttgactg	gatgggcaag	aacctctact	120
gggccgacac	tgggaccaac	agaatcgaag	tggcgcggtg	ggacgggcag	ttccggcaag	180
tcctcgtgtg	gagggacttg	gacaaccga	ggtcgctggc	cctggatccc	accaaggggt	240
aagtgtttgc	ctgtc					255

<210> 50

<211> 210

<212> DNA

<213> Homo sapiens

<400> 50

gtgccttcca	gctacatcta	ctggaccgag	tggggcggca	agccgaggat	cgtgcggggc	60
ttcatggacg	ggaccaactg	catgacgctg	gtggacaagg	tgggccgggc	caacgacctc	120
accattgact	acgctgacca	gcgcctctac	tggaccgacc	tggacaccaa	catgatcgag	180
tcgtccaaca	tgctgggtga	gggccgggct				210

<210> 51

<211> 352

<212> DNA

<213> Homo sapiens

<400> 51

gtgttcatgc	aggtcaggag	cgggtcgtga	ttgccgacga	tctcccgcac	ccgttcggtc	60
tgacgcagta	cagcgattat	atctactgga	cagactggaa	tctgcacagc	attgagcggg	120
ccgacaagac	tagcggccgg	aaccgcaccc	tcatccaggg	ccacctggac	ttcgtgatgg	180
acatcctggt	gttcactcc	tcccgccagg	atggcctcaa	tgactgtatg	cacaacaacg	240
ggcagtgtgg	gcagctgtgc	cttgccatcc	ccggcgggca	ccgctgcggc	tgcgcctcac	300

actacaccct ggaccccage agccgcaact gcagccgtaa gtgcctcatg gt 352

<210> 52

<211> 225

<212> DNA

<213> Homo sapiens

<400> 52

gcctcctcta	cgcccaccac	cttcttgetg	ttcagccaga	aatctgccat	cagtcggatg	60
atcccggacg	accagcacag	cccggatctc	atcctgcccc	tgcattggact	gaggaacgtc	120
aaagccatcg	actatgaccc	actggacaag	ttcatctact	gggtggatgg	gcgccagaac	180
atcaagcgag	ccaaggacga	cgggacccag	gcaggtgccc	tgtgg		225

<210> 53

<211> 235

<212> DNA

<213> Homo sapiens

<400> 53

ctttgtctta	cagccctttg	ttttgacctc	tctgagccaa	ggccaaaacc	cagacaggca	60
gccccacgac	ctcagcatcg	acatctacag	cgggacactg	ttctggacgt	gcgaggccac	120
caataccatc	aacgtccaca	ggctgagcgg	ggaagccatg	gggtgggtgc	tgcgtgggga	180
ccgcgacaag	cccagggcca	tcgtcgtcaa	cgcggagcga	gggtaggagg	ccaac	235

<210> 54

<211> 218

<212> DNA

<213> Homo sapiens

<400> 54

ccaccctccc	gcaggtacct	gtacttcacc	aacatgcagg	accgggcagc	caagatcgaa	60
cgcgcagccc	tggacggcac	cgagcgcgag	gtcctcttca	ccaccggcct	catccgccct	120
gtggcccttg	tgggtggaaa	cacactgggc	aagctgttct	gggtggacgc	ggacctgaag	180
cgcattgaga	gctgtgacct	gtcaggtacg	cgcgcccg			218

<210> 55

<211> 234

<212> DNA

<213> Homo sapiens

<400> 55

ggctgcttgc	aggggccaac	cgctgaccc	tggaggacgc	caacatcgtg	cagcctctgg	60
gcctgacctat	ccttggcaag	catctctact	ggatcgaccg	ccagcagcag	atgatcgagc	120
gtgtggagaa	gaccaccggg	gacaagcgga	ctcgcatcca	gggccgtgtc	gccacctca	180
ctggcatcca	tgcagtggag	gaagtcagcc	tggaggagtt	ctgtacgtgg	gggc	234

<210> 56

<211> 157

<212> DNA

<213> Homo sapiens

<400> 56

ttgtctttgc	agcagccac	ccatgtgccc	gtgacaatgg	tggctgctcc	cacatctgta	60
ttgccaaagg	tgatgggaca	ccacggtgct	catgcccgat	ccacctcgtg	ctcctgcaga	120

acctgctgac ctgtggaggt aggtgtgacc taggtgc 157

<210> 57

<211> 272

<212> DNA

<213> Homo sapiens

<400> 57

gttctcctct	gtccctcccc	cagagccgcc	cacctgctcc	cgggaccagt	ttgcatgtgc	60
cacaggggag	atcgactgta	tccccggggc	ctggcgctgt	gacggctttc	ccgagtgcga	120
tgaccagagc	gacgaggagg	gctgccccgt	gtgctccgcc	gcccagttcc	cctgcgcgcg	180
gggtcagtgt	gtggacctgc	gcctgcgctg	cgacggcgag	gcagactgtc	aggaccgctc	240
agacgaggtg	gactgtgacg	gtgaggccct	cc			272

<210> 58

<211> 134

<212> DNA

<213> Homo sapiens

<400> 58

tctccttgca	gccatctgcc	tgcccaacca	gttccgggtg	gcgagcggcc	agtgtgtcct	60
catcaaacag	cagtgcgact	ccttccccga	ctgtatcgac	ggctccgacg	agctcatgtg	120
tggtgagcca	gctt					134

<210> 59

<211> 274

<212> DNA

<213> Homo sapiens

<400> 59

gtttgtctct	ggcagaaatc	accaagccgc	cctcagacga	cagcccggcc	cacagcagtg	60
ccatcgggcc	cgtcattggc	atcatcctct	ctctcttcgt	catgggtggg	gtctattttg	120
tgtgccagcg	cgtgggtgtg	cagcgctatg	cggggggccaa	cggggcccttc	ccgcacgagt	180
atgtcagcgg	gaccccgccac	gtgcccctca	atttcatagc	cccgggcggg	tcccagcatg	240
gccccttcac	aggtaaggag	cctgagatat	ggaa			274

<210> 60

<211> 164

<212> DNA

<213> Homo sapiens

<400> 60

cttccctgcc	aggcatcgca	tgcgaaaagt	ccatgatgag	ctccgtgagc	ctgatggggg	60
gccggggcgg	gggtgccctc	tacgaccgga	accacgtcac	aggggcctcg	tccagcagct	120
cgtccagcac	gaaggccacg	ctgtaccgcg	cgggtgagggg	cggg		164

<210> 61

<211> 130

<212> DNA

<213> Homo sapiens

<400> 61

ttggctctcc	tcagatcctg	aaccgcgcgc	cctccccggc	cacggacccc	tccctgtaca	60
acatggacat	gttctactct	tcaaacattc	cggccactgc	gagaccgtac	aggtaggaca	120

tccccctgcag

130

<210> 62

<211> 496

<212> DNA

<213> Homo sapiens

<400> 62

tcaaacattc	cggccactgc	gagaccgtac	agggccctaca	tcattcgagg	aatggcgccc	60
ccgacgacgc	cctgcagcac	cgacgtgtgt	gacagcgact	acagcgccag	ccgctggaag	120
gccagcaagt	actacctgga	tttgaactcg	gactcagacc	cctatccacc	cccacccacg	180
ccccacagcc	agtacctgtc	ggcggaggac	agctgcccg	cctcgcccg	caccgagagg	240
agctacttcc	atctcttccc	gccccctccg	tccccctgca	cggactcatc	ctgacctcgg	300
ccggggccact	ctggcttctc	tgtgcccctg	taaatagttt	taaatatgaa	caaagaaaaa	360
aatatatttt	atgattttaa	aaataaatat	aattgggatt	ttaaaaacat	gagaaatgtg	420
aactgtgatg	gggtgggcag	ggctgggaga	actttgtaca	gtggagaaat	atttataaac	480
ttaattttgt	aaaaca					496

<210> 63

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 63

ttttgggtac	acaattcagt	cg	22
------------	------------	----	----

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 64

aaaactgtgg	gtgcttctgg	20
------------	------------	----

<210> 65

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 65

gtgattgagc	caatcctgag	a	21
------------	------------	---	----

<210> 66

<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 66
tgagccaaat aaacccttc t 21

<210> 67
<211> 20
<212> DNA
<213> Homo sapiens

<400> 67
ctggactacg tggccttctc 20

<210> 68
<211> 19
<212> DNA
<213> Homo sapiens

<400> 68
ttcagaagca cttggctgg 19

<210> 69
<211> 20
<212> DNA
<213> Homo sapiens

<400> 69
ctcagtgccca tgaagatgga 20

<210> 70
<211> 21
<212> DNA
<213> Homo sapiens

<400> 70
caagatcact cgatctccag g 21

<210> 71
<211> 20
<212> DNA
<213> Homo sapiens

<400> 71
gtttcaggag actcagagtc 20

<210> 72
<211> 20
<212> DNA
<213> Homo sapiens

<400> 72
ttctgcaggt tgctgttgag 20

<210> 73
<211> 20
<212> DNA
<213> Homo sapiens

<400> 73
ttattgtgat ttcccgtggc 20

<210> 74
<211> 21
<212> DNA
<213> Homo sapiens

<400> 74
gccctctgtc ctgacttcag g 21

<210> 75
<211> 20
<212> DNA
<213> Homo sapiens

<400> 75
gagaaagaaa taaggggacc 20

<210> 76
<211> 20
<212> DNA
<213> Homo sapiens

<400> 76
tgctttgtaa agcactgaga 20

<210> 77
<211> 24
<212> DNA
<213> Homo sapiens

<400> 77
gaagtacggg cagttcagtg gcct 24

<210> 78
<211> 25
<212> DNA
<213> Homo sapiens

<400> 78
atacaccaag gtccatgttc cccgt 25

<210> 79
<211> 25
<212> DNA

<213> Homo sapiens

<400> 79
agcctgggcc acagcgtgag actac 25

<210> 80
<211> 25
<212> DNA
<213> Homo sapiens

<400> 80
tcccggagct tgcacacccg cttca 25

<210> 81
<211> 20
<212> DNA
<213> Homo sapiens

<400> 81
catgtgccca cctcattcat 20

<210> 82
<211> 20
<212> DNA
<213> Homo sapiens

<400> 82
caagattctg tagcttctgg 20

<210> 83
<211> 20
<212> DNA
<213> Homo sapiens

<400> 83
cagagaagtc aagggacttg 20

<210> 84
<211> 20
<212> DNA
<213> Homo sapiens

<400> 84
atcctctcac atcccacact 20

<210> 85
<211> 20
<212> DNA
<213> Homo sapiens

<400> 85
caaggctaaa agacgaaaaa 20

<210> 86

<211> 20
<212> DNA
<213> Homo sapiens

<400> 86
tcaggagcat ttcatctttt 20

<210> 87
<211> 19
<212> DNA
<213> Homo sapiens

<400> 87
aagtcgaggc tgcaaggag 19

<210> 88
<211> 20
<212> DNA
<213> Homo sapiens

<400> 88
gccctgtgtt cctttcagta 20

<210> 89
<211> 19
<212> DNA
<213> Homo sapiens

<400> 89
aaggtgtgag gatcactgg 19

<210> 90
<211> 17
<212> DNA
<213> Homo sapiens

<400> 90
agctcatggg ggctatt 17

<210> 91
<211> 20
<212> DNA
<213> Homo sapiens

<400> 91
gcttctccga gtgtatcaac 20

<210> 92
<211> 20
<212> DNA
<213> Homo sapiens

<400> 92
atggcagagg acttagaaca 20

<210> 93
<211> 24
<212> DNA
<213> Homo sapiens

<400> 93
gatcagcgaa cttcctctcg gctc

24

<210> 94
<211> 24
<212> DNA
<213> Homo sapiens

<400> 94
tccacattga ggactgtggg aacg

24

<210> 95
<211> 20
<212> DNA
<213> Homo sapiens

<400> 95
gctaatacaca gtctaaccga

20

<210> 96
<211> 19
<212> DNA
<213> Homo sapiens

<400> 96
ttgcactgtc ttggatgca

19

<210> 97
<211> 25
<212> DNA
<213> Homo sapiens

<400> 97
gcacagctgt agtgggggttc taggc

25

<210> 98
<211> 25
<212> DNA
<213> Homo sapiens

<400> 98
caggcgcaaa ggacatgcac acggc

25

<210> 99
<211> 25
<212> DNA
<213> Homo sapiens

<400> 99

caccgatgag tgcacgttca aggag	25
<210> 100	
<211> 25	
<212> DNA	
<213> Homo sapiens	
<400> 100	
cagacagaga tgctccacgc catac	25
<210> 101	
<211> 19	
<212> DNA	
<213> Homo sapiens	
<400> 101	
tttctgggtg tgtctgaat	19
<210> 102	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 102	
acacagttgc tctaaagggt	20
<210> 103	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 103	
catttgggaa atccagaaga	20
<210> 104	
<211> 25	
<212> DNA	
<213> Homo sapiens	
<400> 104	
taggtgtctt attttttggt gcttc	25
<210> 105	
<211> 25	
<212> DNA	
<213> Homo sapiens	
<400> 105	
gacataccat gaacactata agagg	25
<210> 106	
<211> 20	
<212> DNA	
<213> Homo sapiens	

<400> 106
caaccatac cagggataag 20

<210> 107
<211> 21
<212> DNA
<213> Homo sapiens

<400> 107
gaacaagagg ggtaagttgg c 21

<210> 108
<211> 22
<212> DNA
<213> Homo sapiens

<400> 108
tgaggacaca gatactgatg gg 22

<210> 109
<211> 25
<212> DNA
<213> Homo sapiens

<400> 109
gaagtgttcc ctcttaaatt ctttg 25

<210> 110
<211> 25
<212> DNA
<213> Homo sapiens

<400> 110
gaactatatt gtagttagtg aggag 25

<210> 111
<211> 18
<212> DNA
<213> Homo sapiens

<400> 111
cctgtaaccc ccagtccc 18

<210> 112
<211> 22
<212> DNA
<213> Homo sapiens

<400> 112
tcttgcttcc taagtttctc gg 22

<210> 113
<211> 21
<212> DNA

<213> Homo sapiens

<400> 113

actccatcca cctcatcact g

21

<210> 114

<211> 20

<212> DNA

<213> Homo sapiens

<400> 114

tgctgtttgc ctcactctgac

20

<210> 115

<211> 20

<212> DNA

<213> Homo sapiens

<400> 115

gtggacaggc atagctgagg

20

<210> 116

<211> 21

<212> DNA

<213> Homo sapiens

<400> 116

tgttcactct tctgcctgca g

21

<210> 117

<211> 20

<212> DNA

<213> Homo sapiens

<400> 117

agctggactc tcacagaatg

20

<210> 118

<211> 20

<212> DNA

<213> Homo sapiens

<400> 118

caagaggctg gtagaagggtg

20

<210> 119

<211> 24

<212> DNA

<213> Homo sapiens

<400> 119

gactccagtc tgggcaataa aagc

24

<210> 120

<211> 22
<212> DNA
<213> Homo sapiens

<400> 120
ggtggcagca tgacctctaa ag

22

<210> 121
<211> 16
<212> DNA
<213> Homo sapiens

<400> 121
caggcccagt ctcttg

16

<210> 122
<211> 18
<212> DNA
<213> Homo sapiens

<400> 122
cgtgtccaga tgaaagtg

18

<210> 123
<211> 18
<212> DNA
<213> Homo sapiens

<400> 123
acctcacggt gtaatccc

18

<210> 124
<211> 18
<212> DNA
<213> Homo sapiens

<400> 124
cttgaagccc atctttgc

18

<210> 125
<211> 23
<212> DNA
<213> Homo sapiens

<400> 125
tatttgcaaa gcttgagact tct

23

<210> 126
<211> 20
<212> DNA
<213> Homo sapiens

<400> 126
aatcactgtg ctttggtgcc

20

<210> 127
<211> 20
<212> DNA
<213> Homo sapiens

<400> 127
actttattgt cagcgtgggc 20

<210> 128
<211> 18
<212> DNA
<213> Homo sapiens

<400> 128
actccctcga tggcttcc 18

<210> 129
<211> 18
<212> DNA
<213> Homo sapiens

<400> 129
gagcagggga gagaaggc 18

<210> 130
<211> 20
<212> DNA
<213> Homo sapiens

<400> 130
cccactggct tgttttattg 20

<210> 131
<211> 25
<212> DNA
<213> Homo sapiens

<400> 131
agccacttta ttgttatattt gatgc 25

<210> 132
<211> 25
<212> DNA
<213> Homo sapiens

<400> 132
aagagtgaac aaaagcaaac atacc 25

<210> 133
<211> 18
<212> DNA
<213> Homo sapiens

<400> 133

gtggagtgtg ggattggg 18

<210> 134
<211> 25
<212> DNA
<213> Homo sapiens

<400> 134
tactgttctt gataagtatg tcggc 25

<210> 135
<211> 25
<212> DNA
<213> Homo sapiens

<400> 135
atgcttttgc atgattctaa ttatt 25

<210> 136
<211> 21
<212> DNA
<213> Homo sapiens

<400> 136
tcccccaaaa gaatgtaaag g 21

<210> 137
<211> 20
<212> DNA
<213> Homo sapiens

<400> 137
ctggctcttcc ttgtgtgctg 20

<210> 138
<211> 18
<212> DNA
<213> Homo sapiens

<400> 138
atcacccagg ccagggat 18

<210> 139
<211> 24
<212> DNA
<213> Homo sapiens

<400> 139
tcagaagcag aactgttttt aaca 24

<210> 140
<211> 22
<212> DNA
<213> Homo sapiens

<400> 140
cctgcttgaa agttctagag cc 22

<210> 141
<211> 21
<212> DNA
<213> Homo sapiens

<400> 141
caagcccggg ttttattgaa a 21

<210> 142
<211> 18
<212> DNA
<213> Homo sapiens

<400> 142
gatgccagga ccatggac 18

<210> 143
<211> 24
<212> DNA
<213> Homo sapiens

<400> 143
gcatatagaa acaatttatt gccg 24

<210> 144
<211> 20
<212> DNA
<213> Homo sapiens

<400> 144
ctctgaagca gggaccagag 20

<210> 145
<211> 19
<212> DNA
<213> Homo sapiens

<400> 145
ctaccacacc acaccaggc 19

<210> 146
<211> 18
<212> DNA
<213> Homo sapiens

<400> 146
caagcgaaag ctgccttc 18

<210> 147
<211> 23
<212> DNA

<213> Homo sapiens

<400> 147

gttgtcttga cttcaggtct gtc

23

<210> 148

<211> 24

<212> DNA

<213> Homo sapiens

<400> 148

ttttccttca acaatcacta ctcc

24

<210> 149

<211> 20

<212> DNA

<213> Homo sapiens

<400> 149

gcgtggggat atagaggtca

20

<210> 150

<211> 20

<212> DNA

<213> Homo sapiens

<400> 150

tacgtggcca agaagctagg

20

<210> 151

<211> 24

<212> DNA

<213> Homo sapiens

<400> 151

taatatatcc ccagtctaag gcat

24

<210> 152

<211> 18

<212> DNA

<213> Homo sapiens

<400> 152

agcttgcaga tggagccc

18

<210> 153

<211> 25

<212> DNA

<213> Homo sapiens

<400> 153

tggtttttaa cctttaatga gaaaa

25

<210> 154

<211> 23
<212> DNA
<213> Homo sapiens

<400> 154
tggtgatcta taccctgttt ccg 23

<210> 155
<211> 25
<212> DNA
<213> Homo sapiens

<400> 155
aattatttaa aagagaggaa aggca 25

<210> 156
<211> 20
<212> DNA
<213> Homo sapiens

<400> 156
tggctgtgaa cttcctctga 20

<210> 157
<211> 25
<212> DNA
<213> Homo sapiens

<400> 157
ggttacagaa aaacatttga gagat 25

<210> 158
<211> 22
<212> DNA
<213> Homo sapiens

<400> 158
tgagctttag ttcccttctc tg 22

<210> 159
<211> 23
<212> DNA
<213> Homo sapiens

<400> 159
ttgaaaaacc atttatttca ccg 23

<210> 160
<211> 18
<212> DNA
<213> Homo sapiens

<400> 160
tctgcggctg ttggattt 18

<210> 161
<211> 23
<212> DNA
<213> Homo sapiens

<400> 161
ttgaaaaacc atttatttca ccg 23

<210> 162
<211> 20
<212> DNA
<213> Homo sapiens

<400> 162
tgttctcttc tcccagcagg 20

<210> 163
<211> 23
<212> DNA
<213> Homo sapiens

<400> 163
ctttattgaa aacattgagt gca 23

<210> 164
<211> 20
<212> DNA
<213> Homo sapiens

<400> 164
ttgtcaaatt ccccccaaaa 20

<210> 165
<211> 16
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 12
<223> n = A,T,C or G

<400> 165
aaaccacgac cnccaa 16

<210> 166
<211> 20
<212> DNA
<213> Homo sapiens

<400> 166
ccctggaaag gtaagatgct 20

<210> 167

<211> 23
<212> DNA
<213> Homo sapiens

<400> 167
cttttggtag agacaagggtc tca 23

<210> 168
<211> 25
<212> DNA
<213> Homo sapiens

<400> 168
tatctgtctg tagtgcttca aatgt 25

<210> 169
<211> 19
<212> DNA
<213> Homo sapiens

<400> 169
gacgaagggtg attcagggc 19

<210> 170
<211> 19
<212> DNA
<213> Homo sapiens

<400> 170
actgaagaac tcttgcct 19

<210> 171
<211> 25
<212> DNA
<213> Homo sapiens

<400> 171
cagataaaag agtcactatg gctca 25

<210> 172
<211> 20
<212> DNA
<213> Homo sapiens

<400> 172
cacttctccc actttgtccc 20

<210> 173
<211> 25
<212> DNA
<213> Homo sapiens

<400> 173
ttattgataa gcattagtga acccc 25

<210> 174
<211> 20
<212> DNA
<213> Homo sapiens

<400> 174
tggcaagtta ggcacagtca 20

<210> 175
<211> 21
<212> DNA
<213> Homo sapiens

<400> 175
ctatgcccag agatgaacag g 21

<210> 176
<211> 20
<212> DNA
<213> Homo sapiens

<400> 176
tccactaagg gctatgtcgc 20

<210> 177
<211> 24
<212> DNA
<213> Homo sapiens

<400> 177
gccagcttta ttgagtaaac ttcc 24

<210> 178
<211> 22
<212> DNA
<213> Homo sapiens

<400> 178
cactggagac tacaagtggg gg 22

<210> 179
<211> 20
<212> DNA
<213> Homo sapiens

<400> 179
catcccaacc atcactcagt 20

<210> 180
<211> 22
<212> DNA
<213> Homo sapiens

<400> 180

ggggactagc ttacagattt ga 22

<210> 181
<211> 23
<212> DNA
<213> Homo sapiens

<400> 181
agactacatt ttggaaccag tgg 23

<210> 182
<211> 23
<212> DNA
<213> Homo sapiens

<400> 182
tgaaaggata tttatagcct gga 23

<210> 183
<211> 20
<212> DNA
<213> Homo sapiens

<400> 183
gaaggttttg tccctcgatc 20

<210> 184
<211> 20
<212> DNA
<213> Homo sapiens

<400> 184
tgagggttgg gaagatcata 20

<210> 185
<211> 18
<212> DNA
<213> Homo sapiens

<400> 185
ccttcatagc cacacccg 18

<210> 186
<211> 21
<212> DNA
<213> Homo sapiens

<400> 186
cagctaactg ttgacatgcc a 21

<210> 187
<211> 25
<212> DNA
<213> Homo sapiens

<400> 187
tcctttactgt gcttacaact ttcct 25

<210> 188
<211> 20
<212> DNA
<213> Homo sapiens

<400> 188
caacagtgcg gtcggtatcg 20

<210> 189
<211> 18
<212> DNA
<213> Homo sapiens

<400> 189
agatcagcaa gcagatag 18

<210> 190
<211> 18
<212> DNA
<213> Homo sapiens

<400> 190
cattccacat ggatagac 18

<210> 191
<211> 24
<212> DNA
<213> Homo sapiens

<400> 191
catacctatg aggtgtgcta cagg 24

<210> 192
<211> 21
<212> DNA
<213> Homo sapiens

<400> 192
gcattttctc atcatccttg c 21

<210> 193
<211> 20
<212> DNA
<213> Homo sapiens

<400> 193
ttacagccac caaggtttcc 20

<210> 194
<211> 20
<212> DNA

<213> Homo sapiens

<400> 194

aggtgtgtgt gccaggttga

20

<210> 195

<211> 25

<212> DNA

<213> Homo sapiens

<400> 195

cactgttatc tcattaactg tgagg

25

<210> 196

<211> 21

<212> DNA

<213> Homo sapiens

<400> 196

tttgattttg tgtctcccaa a

21

<210> 197

<211> 20

<212> DNA

<213> Homo sapiens

<400> 197

ccccactccc acttttattt

20

<210> 198

<211> 24

<212> DNA

<213> Homo sapiens

<400> 198

ccagtcacct ttactagtcc ttg

24

<210> 199

<211> 20

<212> DNA

<213> Homo sapiens

<400> 199

aggacacagc ctgcatctag

20

<210> 200

<211> 20

<212> DNA

<213> Homo sapiens

<400> 200

accaggcatt gcactaaaag

20

<210> 201

<211> 22
<212> DNA
<213> Homo sapiens

<400> 201
gatgggtcac actaacctgt ca 22

<210> 202
<211> 24
<212> DNA
<213> Homo sapiens

<400> 202
acatttatat ttggacatgc aacc 24

<210> 203
<211> 22
<212> DNA
<213> Homo sapiens

<400> 203
agcatcttta atgtgtcagg ca 22

<210> 204
<211> 18
<212> DNA
<213> Homo sapiens

<400> 204
atgtgctggg ctggaaag 18

<210> 205
<211> 20
<212> DNA
<213> Homo sapiens

<400> 205
tcacattcaa aaatcggcaa 20

<210> 206
<211> 18
<212> DNA
<213> Homo sapiens

<400> 206
ctgcctgtgt ggtgtcgc 18

<210> 207
<211> 25
<212> DNA
<213> Homo sapiens

<400> 207
tgttttatatt ctcaagtacaa agcca 25

<210> 208
<211> 19
<212> DNA
<213> Homo sapiens

<400> 208
gacctcctgt gacaccacg 19

<210> 209
<211> 25
<212> DNA
<213> Homo sapiens

<400> 209
ccaccaaatt atttatagtt ctgcg 25

<210> 210
<211> 23
<212> DNA
<213> Homo sapiens

<400> 210
gtaagattct ccactggtgc acc 23

<210> 211
<211> 20
<212> DNA
<213> Homo sapiens

<400> 211
cctataatgg gctggaccaa 20

<210> 212
<211> 21
<212> DNA
<213> Homo sapiens

<400> 212
actcctcatg tgaagtcacc g 21

<210> 213
<211> 20
<212> DNA
<213> Homo sapiens

<400> 213
cagtgtgcac gttttcattt 20

<210> 214
<211> 20
<212> DNA
<213> Homo sapiens

<400> 214

cagcatcttc agcacttacc 20

<210> 215
<211> 25
<212> DNA
<213> Homo sapiens

<400> 215
ctgcatTTat tatgagaatc aacag 25

<210> 216
<211> 20
<212> DNA
<213> Homo sapiens

<400> 216
tgctgctggg agtcagagtc 20

<210> 217
<211> 23
<212> DNA
<213> Homo sapiens

<400> 217
cagggcactg agatacactt acc 23

<210> 218
<211> 21
<212> DNA
<213> Homo sapiens

<400> 218
aaggatcaag ccaggcattt g 21

<210> 219
<211> 20
<212> DNA
<213> Homo sapiens

<400> 219
acacatctct tctgtgcccc 20

<210> 220
<211> 18
<212> DNA
<213> Homo sapiens

<400> 220
tgaaccctgg aggcagag 18

<210> 221
<211> 19
<212> DNA
<213> Homo sapiens

<400> 221
cattccccag ttgcagac 19

<210> 222
<211> 19
<212> DNA
<213> Homo sapiens

<400> 222
gtgctgggat tacaggtgt 19

<210> 223
<211> 20
<212> DNA
<213> Homo sapiens

<400> 223
gcagagaagt cctgttagcc 20

<210> 224
<211> 20
<212> DNA
<213> Homo sapiens

<400> 224
ccatgctaga gaagcacaac 20

<210> 225
<211> 20
<212> DNA
<213> Homo sapiens

<400> 225
agtgtggggc aggacctctg 20

<210> 226
<211> 21
<212> DNA
<213> Homo sapiens

<400> 226
cagacagata gccctggggt c 21

<210> 227
<211> 20
<212> DNA
<213> Homo sapiens

<400> 227
tccctcatcc ccttgtctgt 20

<210> 228
<211> 19
<212> DNA

<213> Homo sapiens

<400> 228
agccccctg gggataatc 19

<210> 229
<211> 19
<212> DNA
<213> Homo sapiens

<400> 229
gatgcttacc taccacggc 19

<210> 230
<211> 21
<212> DNA
<213> Homo sapiens

<400> 230
aggattccta tctgggctat g 21

<210> 231
<211> 20
<212> DNA
<213> Homo sapiens

<400> 231
tggcagacca tgctccgcct 20

<210> 232
<211> 20
<212> DNA
<213> Homo sapiens

<400> 232
gagaaggccg ggaggctctg 20

<210> 233
<211> 25
<212> DNA
<213> Homo sapiens

<400> 233
ctccatcaca accagatttg aggct 25

<210> 234
<211> 22
<212> DNA
<213> Homo sapiens

<400> 234
gggtgtgagc tgctgctgaa gg 22

<210> 235

<211> 25
<212> DNA
<213> Homo sapiens

<400> 235
agtgggaaac ctcaggtagc tccccg 25

<210> 236
<211> 25
<212> DNA
<213> Homo sapiens

<400> 236
cagtttggct cagacatatg ggggc 25

<210> 237
<211> 20
<212> DNA
<213> Homo sapiens

<400> 237
cattagtagt ggggggacag 20

<210> 238
<211> 21
<212> DNA
<213> Homo sapiens

<400> 238
caaagcgaca gtgagttagg g 21

<210> 239
<211> 20
<212> DNA
<213> Homo sapiens

<400> 239
ggagtagacc atgattactg 20

<210> 240
<211> 19
<212> DNA
<213> Homo sapiens

<400> 240
catggtctat ttattctcg 19

<210> 241
<211> 22
<212> DNA
<213> Homo sapiens

<400> 241
cgccctggat cctcacacta ca 22

<210> 242
<211> 21
<212> DNA
<213> Homo sapiens

<400> 242
gggcatcagg ggatgggtag a 21

<210> 243
<211> 23
<212> DNA
<213> Homo sapiens

<400> 243
gctcctatct gtgttttgaa tgg 23

<210> 244
<211> 22
<212> DNA
<213> Homo sapiens

<400> 244
ccgtggcata gataagtaaa cg 22

<210> 245
<211> 22
<212> DNA
<213> Homo sapiens

<400> 245
cttggagcgc tatgaggagg gc 22

<210> 246
<211> 23
<212> DNA
<213> Homo sapiens

<400> 246
atggcaactg accttccgtc ctg 23

<210> 247
<211> 16
<212> DNA
<213> Homo sapiens

<400> 247
ttggagtcac aggggc 16

<210> 248
<211> 17
<212> DNA
<213> Homo sapiens

<400> 248

cagcactatc cttgggg 17

<210> 249
<211> 21
<212> DNA
<213> Homo sapiens

<400> 249
aacaaagctg cttagcacct g 21

<210> 250
<211> 20
<212> DNA
<213> Homo sapiens

<400> 250
gatgaggacc aactggtgac 20

<210> 251
<211> 20
<212> DNA
<213> Homo sapiens

<400> 251
ttttccaata atgtgacttc 20

<210> 252
<211> 20
<212> DNA
<213> Homo sapiens

<400> 252
caatcccaac cgtaacaggc 20

<210> 253
<211> 19
<212> DNA
<213> Homo sapiens

<400> 253
cttgatctcg cccaggaac 19

<210> 254
<211> 20
<212> DNA
<213> Homo sapiens

<400> 254
gctcgctgaa ggatgaagac 20

<210> 255
<211> 17
<212> DNA
<213> Homo sapiens

<400> 255
gaatcgcttg aaccag 17

<210> 256
<211> 17
<212> DNA
<213> Homo sapiens

<400> 256
ccaggtggtc ttaacgg 17

<210> 257
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 8
<223> n = A,T,C or G

<400> 257
gaacgttntt catgtaggcg t 21

<210> 258
<211> 16
<212> DNA
<213> Homo sapiens

<400> 258
taatggtcgc tgtccc 16

<210> 259
<211> 22
<212> DNA
<213> Homo sapiens

<400> 259
agggaaaatg gtatgtgggg ag 22

<210> 260
<211> 18
<212> DNA
<213> Homo sapiens

<400> 260
gcagtgtgtg aaggcagg 18

<210> 261
<211> 24
<212> DNA
<213> Homo sapiens

<400> 261
agtggacaaa atgaggaaaa cagg 24

<210> 262
<211> 23
<212> DNA
<213> Homo sapiens

<400> 262
ccaacacagt ttgctcacat gcc 23

<210> 263
<211> 20
<212> DNA
<213> Homo sapiens

<400> 263
tgacatcttt gcattatggc 20

<210> 264
<211> 20
<212> DNA
<213> Homo sapiens

<400> 264
agttatccca cctgataccg 20

<210> 265
<211> 20
<212> DNA
<213> Homo sapiens

<400> 265
agctcttgct tctcagtcca 20

<210> 266
<211> 24
<212> DNA
<213> Homo sapiens

<400> 266
caaaagttgt ttctgtgttt gttc 24

<210> 267
<211> 22
<212> DNA
<213> Homo sapiens

<400> 267
gcctctcaaa gtagttggaa cc 22

<210> 268
<211> 23
<212> DNA
<213> Homo sapiens

<400> 268

tgtgtatcca tagtgcaaaa cag 23

<210> 269
<211> 19
<212> DNA
<213> Homo sapiens

<400> 269
ctcaaggcca ggcattcact 19

<210> 270
<211> 19
<212> DNA
<213> Homo sapiens

<400> 270
ggactcttcc atgccagtg 19

<210> 271
<211> 19
<212> DNA
<213> Homo sapiens

<400> 271
aatgatgatc tcaactctg 19

<210> 272
<211> 19
<212> DNA
<213> Homo sapiens

<400> 272
actgaagaac tcttgcct 19

<210> 273
<211> 23
<212> DNA
<213> Homo sapiens

<400> 273
gacatctggt agtctcataa ttc 23

<210> 274
<211> 18
<212> DNA
<213> Homo sapiens

<400> 274
ggtaacagtg tcttgctt 18

<210> 275
<211> 21
<212> DNA
<213> Homo sapiens

<400> 275
ctatgtacaa aacaggaaga g 21

<210> 276
<211> 19
<212> DNA
<213> Homo sapiens

<400> 276
atcctagttt cctctcctt 19

<210> 277
<211> 23
<212> DNA
<213> Homo sapiens

<400> 277
gtaaatgaga aacagacaaa tga 23

<210> 278
<211> 23
<212> DNA
<213> Homo sapiens

<400> 278
ctattggatg tgatatgtta tgg 23

<210> 279
<211> 22
<212> DNA
<213> Homo sapiens

<400> 279
aagtagaaac aaaatgaggg ac 22

<210> 280
<211> 18
<212> DNA
<213> Homo sapiens

<400> 280
cctaccccaa ggtaacag 18

<210> 281
<211> 22
<212> DNA
<213> Homo sapiens

<400> 281
acttcctata aatggaggtg ag 22

<210> 282
<211> 18
<212> DNA

<213> Homo sapiens

<400> 282
gaggagcttc aagaggaa 18

<210> 283
<211> 23
<212> DNA
<213> Homo sapiens

<400> 283
catactccta gactcaagga atc 23

<210> 284
<211> 23
<212> DNA
<213> Homo sapiens

<400> 284
gaatgatgta catgaattct ttg 23

<210> 285
<211> 20
<212> DNA
<213> Homo sapiens

<400> 285
gtgttgagga gaaaagcact 20

<210> 286
<211> 19
<212> DNA
<213> Homo sapiens

<400> 286
ctcccagtag tcacattcc 19

<210> 287
<211> 23
<212> DNA
<213> Homo sapiens

<400> 287
caagttacaa ataacttaag ccg 23

<210> 288
<211> 23
<212> DNA
<213> Homo sapiens

<400> 288
caagacccta tctctacaaa aac 23

<210> 289

<211> 24
<212> DNA
<213> Homo sapiens

<400> 289
tttattagaa gtgactcttg gccc 24

<210> 290
<211> 20
<212> DNA
<213> Homo sapiens

<400> 290
gactacctgc cctcagcttg 20

<210> 291
<211> 20
<212> DNA
<213> Homo sapiens

<400> 291
ttctcatgta caaagcggtc 20

<210> 292
<211> 20
<212> DNA
<213> Homo sapiens

<400> 292
ccactggctt ctctcttttt 20

<210> 293
<211> 18
<212> DNA
<213> Homo sapiens

<400> 293
caccagaagg ttgggggtg 18

<210> 294
<211> 21
<212> DNA
<213> Homo sapiens

<400> 294
actattacga catgaacgcg g 21

<210> 295
<211> 18
<212> DNA
<213> Homo sapiens

<400> 295
ctcatgctgg atgacccc 18

<210> 296
<211> 24
<212> DNA
<213> Homo sapiens

<400> 296
ttgcctttct tgaaacttaa ttcc 24

<210> 297
<211> 19
<212> DNA
<213> Homo sapiens

<400> 297
tcacagcctt cagtcaggg 19

<210> 298
<211> 18
<212> DNA
<213> Homo sapiens

<400> 298
acatgctgtg gcaccatg 18

<210> 299
<211> 18
<212> DNA
<213> Homo sapiens

<400> 299
cctgagctac tgccacag 18

<210> 300
<211> 20
<212> DNA
<213> Homo sapiens

<400> 300
ccctgacttg gacagtgtcc 20

<210> 301
<211> 18
<212> DNA
<213> Homo sapiens

<400> 301
tcagagtcac tcctgccc 18

<210> 302
<211> 22
<212> DNA
<213> Homo sapiens

<400> 302

caaattcaag ctcatccaga cc

22

<210> 303

<211> 18

<212> DNA

<213> Homo sapiens

<400> 303

cggcatttca tccaggac

18

<210> 304

<211> 20

<212> DNA

<213> Homo sapiens

<400> 304

ggtgtaggag gtgcgacaat

20

<210> 305

<211> 20

<212> DNA

<213> Homo sapiens

<400> 305

ttccatttat tgagcacctg

20

<210> 306

<211> 20

<212> DNA

<213> Homo sapiens

<400> 306

cttaagccac tgtgttttgg

20

<210> 307

<211> 20

<212> DNA

<213> Homo sapiens

<400> 307

cctcctacac ctgcaaaagc

20

<210> 308

<211> 19

<212> DNA

<213> Homo sapiens

<400> 308

tggaagaacc ccagaggac

19

<210> 309

<211> 23

<212> DNA

<213> Homo sapiens

<400> 309
aaagcacaaa agtaacagca aca 23

<210> 310
<211> 20
<212> DNA
<213> Homo sapiens

<400> 310
gtgtgtgggc cacaatattg 20

<210> 311
<211> 20
<212> DNA
<213> Homo sapiens

<400> 311
agagcacctt tcctcagcac 20

<210> 312
<211> 20
<212> DNA
<213> Homo sapiens

<400> 312
agaatctcat cacaggggcg 20

<210> 313
<211> 25
<212> DNA
<213> Homo sapiens

<400> 313
aaaaaggaca gtgtctaaaa tttga 25

<210> 314
<211> 25
<212> DNA
<213> Homo sapiens

<400> 314
aattgttttt gtttgttttt tgagt 25

<210> 315
<211> 22
<212> DNA
<213> Homo sapiens

<400> 315
gatttaggga gtacaagtgc gg 22

<210> 316
<211> 25
<212> DNA

<213> Homo sapiens

<400> 316

ggggacaaat tatactttat tcagg

25

<210> 317

<211> 22

<212> DNA

<213> Homo sapiens

<400> 317

ccatcatcat attggtgtga cc

22

<210> 318

<211> 18

<212> DNA

<213> Homo sapiens

<400> 318

tggctgccca agaagaag

18

<210> 319

<211> 24

<212> DNA

<213> Homo sapiens

<400> 319

ttaagatgcc attaaactca tgac

24

<210> 320

<211> 20

<212> DNA

<213> Homo sapiens

<400> 320

ccaaggagat gaccaagtgg

20

<210> 321

<211> 22

<212> DNA

<213> Homo sapiens

<400> 321

ccatctcttt tatcagggtt gg

22

<210> 322

<211> 24

<212> DNA

<213> Homo sapiens

<400> 322

ctctgtgcaa gtaagcatct taca

24

<210> 323

<211> 20
<212> DNA
<213> Homo sapiens

<400> 323
cgactgtgta ttttccacag 20

<210> 324
<211> 20
<212> DNA
<213> Homo sapiens

<400> 324
agaagcccat atcaatgcac 20

<210> 325
<211> 23
<212> DNA
<213> Homo sapiens

<400> 325
agcttaaagt aggacaacca tgg 23

<210> 326
<211> 20
<212> DNA
<213> Homo sapiens

<400> 326
ggatgcttca ctccagaaag 20

<210> 327
<211> 21
<212> DNA
<213> Homo sapiens

<400> 327
tgttgtttat ttccacctgc c 21

<210> 328
<211> 18
<212> DNA
<213> Homo sapiens

<400> 328
agagtggctg caggccag 18

<210> 329
<211> 25
<212> DNA
<213> Homo sapiens

<400> 329
tttttttttt tacacgaatt tgagg 25

<210> 330
<211> 23
<212> DNA
<213> Homo sapiens

<400> 330
tgaggaagta aaaacaggtc atc 23

<210> 331
<211> 23
<212> DNA
<213> Homo sapiens

<400> 331
atgaaatctt aagcagaatc cca 23

<210> 332
<211> 20
<212> DNA
<213> Homo sapiens

<400> 332
cacagagtcc cagggctctgt 20

<210> 333
<211> 25
<212> DNA
<213> Homo sapiens

<400> 333
aaaggccttt atttatctct ctctg 25

<210> 334
<211> 18
<212> DNA
<213> Homo sapiens

<400> 334
gcctcagagc tgggtgggt 18

<210> 335
<211> 25
<212> DNA
<213> Homo sapiens

<400> 335
gcttctaagt cttagagtca gctgg 25

<210> 336
<211> 19
<212> DNA
<213> Homo sapiens

<400> 336

agcccacagt cagcctacc 19

<210> 337
<211> 20
<212> DNA
<213> Homo sapiens

<400> 337
ttggttaa at gatgccaga 20

<210> 338
<211> 18
<212> DNA
<213> Homo sapiens

<400> 338
tggtcccact cacatccc 18

<210> 339
<211> 19
<212> DNA
<213> Homo sapiens

<400> 339
acacagcatg cagggagag 19

<210> 340
<211> 19
<212> DNA
<213> Homo sapiens

<400> 340
atccctggtg cttagggtg 19

<210> 341
<211> 21
<212> DNA
<213> Homo sapiens

<400> 341
gatggaagta gctcctctcg g 21

<210> 342
<211> 21
<212> DNA
<213> Homo sapiens

<400> 342
ggaaggccag caagtactac c 21

<210> 343
<211> 18
<212> DNA
<213> Homo sapiens

<400> 343
ccggtgcttg gaaagatg 18

<210> 344
<211> 20
<212> DNA
<213> Homo sapiens

<400> 344
gaagtgtctc tggtggggga 20

<210> 345
<211> 22
<212> DNA
<213> Homo sapiens

<400> 345
ttacaggcat gagtcactac gc 22

<210> 346
<211> 21
<212> DNA
<213> Homo sapiens

<400> 346
accactctca cagcccttac a 21

<210> 347
<211> 18
<212> DNA
<213> Homo sapiens

<400> 347
ccctccctcc acacacac 18

<210> 348
<211> 20
<212> DNA
<213> Homo sapiens

<400> 348
gctcactgaa ctttcagggc 20

<210> 349
<211> 20
<212> DNA
<213> Homo sapiens

<400> 349
agatacgggc aaaacactgg 20

<210> 350
<211> 21
<212> DNA

<213> Homo sapiens

<400> 350
gttgaatata gagcagggcc c 21

<210> 351
<211> 20
<212> DNA
<213> Homo sapiens

<400> 351
ttctgaggtc agggctgtct 20

<210> 352
<211> 21
<212> DNA
<213> Homo sapiens

<400> 352
agcttgga aa atctcgtgtc a 21

<210> 353
<211> 18
<212> DNA
<213> Homo sapiens

<400> 353
actcagtccc tcccaccc 18

<210> 354
<211> 20
<212> DNA
<213> Homo sapiens

<400> 354
tcctctcact ccttcccaga 20

<210> 355
<211> 19
<212> DNA
<213> Homo sapiens

<400> 355
gtgatcacgg ctcaacctg 19

<210> 356
<211> 18
<212> DNA
<213> Homo sapiens

<400> 356
tggaggactg cttgagcc 18

<210> 357

<211> 19
<212> DNA
<213> Homo sapiens

<400> 357
ctgcagctgc ctcagtttc 19

<210> 358
<211> 20
<212> DNA
<213> Homo sapiens

<400> 358
tcaaaaagtgc tggtagacagc 20

<210> 359
<211> 20
<212> DNA
<213> Homo sapiens

<400> 359
atttccagag ccagctcaaa 20

<210> 360
<211> 25
<212> DNA
<213> Homo sapiens

<400> 360
ctttaatgtt gtgatgacac aaagc 25

<210> 361
<211> 20
<212> DNA
<213> Homo sapiens

<400> 361
gatcatgcac tgtagaccac 20

<210> 362
<211> 25
<212> DNA
<213> Homo sapiens

<400> 362
tacatttgaa acatttaaaa cctga 25

<210> 363
<211> 24
<212> DNA
<213> Homo sapiens

<400> 363
aactgagctg taaccagact ggga 24

<210> 364
<211> 22
<212> DNA
<213> Homo sapiens

<400> 364
tggaacagtc tggtcctgat gg 22

<210> 365
<211> 25
<212> DNA
<213> Homo sapiens

<400> 365
ttatcccttt attgtttctc ctttg 25

<210> 366
<211> 24
<212> DNA
<213> Homo sapiens

<400> 366
tggtcacctg tatttattgc tagg 24

<210> 367
<211> 22
<212> DNA
<213> Homo sapiens

<400> 367
tcttcaaagc ctctgcagta cc 22

<210> 368
<211> 22
<212> DNA
<213> Homo sapiens

<400> 368
ctcatctcca acctgtctaa cc 22

<210> 369
<211> 24
<212> DNA
<213> Homo sapiens

<400> 369
gtggctgcag ctaatgtaag acac 24

<210> 370
<211> 24
<212> DNA
<213> Homo sapiens

<400> 370

cagcagagac aatggcgtaa gtcc

24

<210> 371

<211> 20

<212> DNA

<213> Homo sapiens

<400> 371

ctgattgaga accagaacag

20

<210> 372

<211> 20

<212> DNA

<213> Homo sapiens

<400> 372

taaagcccta taacctctcc

20

<210> 373

<211> 20

<212> DNA

<213> Homo sapiens

<400> 373

tagtaagggga ccttcaccag

20

<210> 374

<211> 20

<212> DNA

<213> Homo sapiens

<400> 374

agatgtttgg tatgacttgg

20

<210> 375

<211> 20

<212> DNA

<213> Homo sapiens

<400> 375

gatgattaaa ctctcctggc

20

<210> 376

<211> 20

<212> DNA

<213> Homo sapiens

<400> 376

gagacagcta agcactcatg

20

<210> 377

<211> 19

<212> DNA

<213> Homo sapiens

<400> 377
gaggtggtgg gcacctgta 19

<210> 378
<211> 20
<212> DNA
<213> Homo sapiens

<400> 378
agaggggagg aacacacctt 20

<210> 379
<211> 20
<212> DNA
<213> Homo sapiens

<400> 379
gaccagagtc tgcccagaag 20

<210> 380
<211> 19
<212> DNA
<213> Homo sapiens

<400> 380
tccccagctc tatcccaac 19

<210> 381
<211> 20
<212> DNA
<213> Homo sapiens

<400> 381
ggaggggatgg acaagtctga 20

<210> 382
<211> 20
<212> DNA
<213> Homo sapiens

<400> 382
gtccagctcg ctgactatcc 20

<210> 383
<211> 20
<212> DNA
<213> Homo sapiens

<400> 383
tcaaaacaca gtcattctca 20

<210> 384
<211> 20
<212> DNA

<213> Homo sapiens

<400> 384
gcaaaggctt taccatattg 20

<210> 385
<211> 16
<212> DNA
<213> Homo sapiens

<400> 385
gctcagcacc cccatt 16

<210> 386
<211> 16
<212> DNA
<213> Homo sapiens

<400> 386
tccctgctcg ggaaac 16

<210> 387
<211> 20
<212> DNA
<213> Homo sapiens

<400> 387
gttctccaga gagacagcac 20

<210> 388
<211> 19
<212> DNA
<213> Homo sapiens

<400> 388
gagagcaaca ctattgccc 19

<210> 389
<211> 21
<212> DNA
<213> Homo sapiens

<400> 389
tatagacttc agccctgctg c 21

<210> 390
<211> 20
<212> DNA
<213> Homo sapiens

<400> 390
cctctgtagg atgcagttgg 20

<210> 391

<211> 20
<212> DNA
<213> Homo sapiens

<400> 391
ttgctacgca ctctctact 20

<210> 392
<211> 20
<212> DNA
<213> Homo sapiens

<400> 392
gtgaaggcag gaaatgtgac 20

<210> 393
<211> 20
<212> DNA
<213> Homo sapiens

<400> 393
atcctagacc agaggagccc 20

<210> 394
<211> 20
<212> DNA
<213> Homo sapiens

<400> 394
ctccccctgg tccagttatt 20

<210> 395
<211> 20
<212> DNA
<213> Homo sapiens

<400> 395
aactttcatt tgccaaggga 20

<210> 396
<211> 20
<212> DNA
<213> Homo sapiens

<400> 396
agcagatctg ctcttgcat 20

<210> 397
<211> 20
<212> DNA
<213> Homo sapiens

<400> 397
acagttgtca tcggtaggca 20

<210> 398
<211> 22
<212> DNA
<213> Homo sapiens

<400> 398
aaaagtatga atgggatgga gc 22

<210> 399
<211> 20
<212> DNA
<213> Homo sapiens

<400> 399
gtgcaggtgg cgtttatattt 20

<210> 400
<211> 21
<212> DNA
<213> Homo sapiens

<400> 400
ccctatatct ccgtgtgctc c 21

<210> 401
<211> 21
<212> DNA
<213> Homo sapiens

<400> 401
gctctagtgg gaaacctcag g 21

<210> 402
<211> 20
<212> DNA
<213> Homo sapiens

<400> 402
gaattccagg ctcttgcttg 20

<210> 403
<211> 20
<212> DNA
<213> Homo sapiens

<400> 403
ggtttggtct caaaggcaaa 20

<210> 404
<211> 20
<212> DNA
<213> Homo sapiens

<400> 404

ccagtacatg gtggtcacca	20
<210> 405	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 405	
gctgccttgg aatttctggt	20
<210> 406	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 406	
gtgctgtggt ggggaaag	18
<210> 407	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 407	
attcaagctc atccagaccc	20
<210> 408	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 408	
ggactggccc tttgaaactc	20
<210> 409	
<211> 22	
<212> DNA	
<213> Homo sapiens	
<400> 409	
atattgaccg tgcacaaata cg	22
<210> 410	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 410	
agacctggga aaagtggaga a	21
<210> 411	
<211> 20	
<212> DNA	
<213> Homo sapiens	

<400> 411
attggcagtg gaaaatgctt 20

<210> 412
<211> 25
<212> DNA
<213> Homo sapiens

<400> 412
ttaatctttt gtcaacttcc tgatt 25

<210> 413
<211> 23
<212> DNA
<213> Homo sapiens

<400> 413
tctgtcctcc ttccaccgga agc 23

<210> 414
<211> 29
<212> DNA
<213> Homo sapiens

<400> 414
ggataaagaa actccgctct gctggtaga 29

<210> 415
<211> 25
<212> DNA
<213> Homo sapiens

<400> 415
tcagggcctg tgttgccgca ctctg 25

<210> 416
<211> 25
<212> DNA
<213> Homo sapiens

<400> 416
agcgatgtaa aggggtaccag tgccg 25

<210> 417
<211> 18
<212> DNA
<213> Homo sapiens

<400> 417
aggcatgcaa gcttctta 18

<210> 418
<211> 18
<212> DNA

<213> Homo sapiens

<400> 418
ccgggaggag acatctat 18

<210> 419
<211> 19
<212> DNA
<213> Homo sapiens

<400> 419
tggttaagcac agaaaatgc 19

<210> 420
<211> 18
<212> DNA
<213> Homo sapiens

<400> 420
aatggatggg ggattatt 18

<210> 421
<211> 18
<212> DNA
<213> Homo sapiens

<400> 421
ctggacgtta tgtctgcc 18

<210> 422
<211> 18
<212> DNA
<213> Homo sapiens

<400> 422
agaggcccag tcacagat 18

<210> 423
<211> 19
<212> DNA
<213> Homo sapiens

<400> 423
atcactctga actgccact 19

<210> 424
<211> 20
<212> DNA
<213> Homo sapiens

<400> 424
cccttctggt tttctgtttt 20

<210> 425

<211> 18
<212> DNA
<213> Homo sapiens

<400> 425
caagctttga aggaagag 18

<210> 426
<211> 19
<212> DNA
<213> Homo sapiens

<400> 426
taggacgtta agtgaggac 19

<210> 427
<211> 18
<212> DNA
<213> Homo sapiens

<400> 427
gctctgcagt gggtaaaa 18

<210> 428
<211> 18
<212> DNA
<213> Homo sapiens

<400> 428
actctccaag actgtgcg 18

<210> 429
<211> 18
<212> DNA
<213> Homo sapiens

<400> 429
ccctttctga ggcaagat 18

<210> 430
<211> 18
<212> DNA
<213> Homo sapiens

<400> 430
gaccacctgg gagagaac 18

<210> 431
<211> 18
<212> DNA
<213> Homo sapiens

<400> 431
cgctatgagt cccatctg 18

<210> 432
<211> 18
<212> DNA
<213> Homo sapiens

<400> 432
gatcagctgc aatgaagg

18

<210> 433
<211> 18
<212> DNA
<213> Homo sapiens

<400> 433
ttgagtacac ggggtgac

18

<210> 434
<211> 18
<212> DNA
<213> Homo sapiens

<400> 434
cgcaggactg aaagatga

18

<210> 435
<211> 18
<212> DNA
<213> Homo sapiens

<400> 435
acctgtctcc tctcctgg

18

<210> 436
<211> 18
<212> DNA
<213> Homo sapiens

<400> 436
tgcttttctt ctgtggga

18

<210> 437
<211> 18
<212> DNA
<213> Homo sapiens

<400> 437
atgaccagca agcattgt

18

<210> 438
<211> 18
<212> DNA
<213> Homo sapiens

<400> 438

gtactgggat tacaggcg	18
<210> 439	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 439	
gcagaaggtc ctttggat	18
<210> 440	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 440	
tttgcaggat tcatgctt	18
<210> 441	
<211> 19	
<212> DNA	
<213> Homo sapiens	
<400> 441	
cgacattctt ttctggagg	19
<210> 442	
<211> 19	
<212> DNA	
<213> Homo sapiens	
<400> 442	
acctttgcat gttggtttt	19
<210> 443	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 443	
gcacttttcc ttccttcc	18
<210> 444	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 444	
tgctttgctt tcttctgg	18
<210> 445	
<211> 19	
<212> DNA	
<213> Homo sapiens	

<400> 445 acagctccag agagaagga	19
<210> 446 <211> 19 <212> DNA <213> Homo sapiens	
<400> 446 gcagtcactt gaaaccaga	19
<210> 447 <211> 18 <212> DNA <213> Homo sapiens	
<400> 447 aggcatcaag ctttcctt	18
<210> 448 <211> 18 <212> DNA <213> Homo sapiens	
<400> 448 ggtttagaga accgagcc	18
<210> 449 <211> 18 <212> DNA <213> Homo sapiens	
<400> 449 gtggtgctgc aagttacc	18
<210> 450 <211> 18 <212> DNA <213> Homo sapiens	
<400> 450 ggaatccctt tctttcca	18
<210> 451 <211> 18 <212> DNA <213> Homo sapiens	
<400> 451 gaccatttgt tacgcagc	18
<210> 452 <211> 19 <212> DNA	

<213> Homo sapiens

<400> 452
gatgggtgtg aatgaacaa 19

<210> 453
<211> 19
<212> DNA
<213> Homo sapiens

<400> 453
ctcaagcttc tgttcatgc 19

<210> 454
<211> 18
<212> DNA
<213> Homo sapiens

<400> 454
gctgtgagtg tcttggt 18

<210> 455
<211> 18
<212> DNA
<213> Homo sapiens

<400> 455
tacagaaaac cgcagctc 18

<210> 456
<211> 18
<212> DNA
<213> Homo sapiens

<400> 456
gccaccaaag gaaagatt 18

<210> 457
<211> 18
<212> DNA
<213> Homo sapiens

<400> 457
aaaaggaggg aatcatgg 18

<210> 458
<211> 18
<212> DNA
<213> Homo sapiens

<400> 458
tcacttagca ggaggcag 18

<210> 459

<211> 18
<212> DNA
<213> Homo sapiens

<400> 459
ctgagcatcc gatgagac 18

<210> 460
<211> 18
<212> DNA
<213> Homo sapiens

<400> 460
gtgcaaaatg agcagctt 18

<210> 461
<211> 18
<212> DNA
<213> Homo sapiens

<400> 461
tctaaccctt tactgggc 18

<210> 462
<211> 18
<212> DNA
<213> Homo sapiens

<400> 462
tcctcaaact gggaatga 18

<210> 463
<211> 18
<212> DNA
<213> Homo sapiens

<400> 463
tttacacagg accaggga 18

<210> 464
<211> 18
<212> DNA
<213> Homo sapiens

<400> 464
atctcccca ctcagaag 18

<210> 465
<211> 18
<212> DNA
<213> Homo sapiens

<400> 465
gtccacgggc tttattct 18

<210> 466
<211> 22
<212> DNA
<213> Homo sapiens

<400> 466
tgagcataaa tttcattagc tg 22

<210> 467
<211> 20
<212> DNA
<213> Homo sapiens

<400> 467
ggaagagcaa aataaatcca 20

<210> 468
<211> 19
<212> DNA
<213> Homo sapiens

<400> 468
ggtgcacaga attgttcat 19

<210> 469
<211> 18
<212> DNA
<213> Homo sapiens

<400> 469
agcacgctta tttcatgg 18

<210> 470
<211> 18
<212> DNA
<213> Homo sapiens

<400> 470
gtaacaccag cagggaca 18

<210> 471
<211> 18
<212> DNA
<213> Homo sapiens

<400> 471
tcctgctgca ttatggat 18

<210> 472
<211> 18
<212> DNA
<213> Homo sapiens

<400> 472

gggggtgaga agtaggaa	18
<210> 473	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 473	
atggggatta aatacggg	18
<210> 474	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 474	
agctagcatt gggctctt	18
<210> 475	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 475	
ctgaggagaa gaggctgg	18
<210> 476	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 476	
cgcccttaca ggcaagta	18
<210> 477	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 477	
aggatgcttg ctagggtt	18
<210> 478	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 478	
cacaagtgtc tggaaggc	18
<210> 479	
<211> 18	
<212> DNA	
<213> Homo sapiens	

<400> 479
ggtctcagga gcccttta 18

<210> 480
<211> 21
<212> DNA
<213> Homo sapiens

<400> 480
acatgccact cttctcacta a 21

<210> 481
<211> 18
<212> DNA
<213> Homo sapiens

<400> 481
acttaaccaa ggatgggg 18

<210> 482
<211> 18
<212> DNA
<213> Homo sapiens

<400> 482
caaccacga gcataaga 18

<210> 483
<211> 18
<212> DNA
<213> Homo sapiens .

<400> 483
taggctctgc actcttgg 18

<210> 484
<211> 18
<212> DNA
<213> Homo sapiens

<400> 484
accacaggag tctctctc 18

<210> 485
<211> 18
<212> DNA
<213> Homo sapiens

<400> 485
taaaggcggg gaagtgg 18

<210> 486
<211> 18
<212> DNA

<213> Homo sapiens

<400> 486

ctaccgctct cctaggct 18

<210> 487

<211> 18

<212> DNA

<213> Homo sapiens

<400> 487

tggggccaga taattctt 18

<210> 488

<211> 18

<212> DNA

<213> Homo sapiens

<400> 488

ctggtgtttg gtggtggt 18

<210> 489

<211> 18

<212> DNA

<213> Homo sapiens

<400> 489

aaggaagagg tcaccagg 18

<210> 490

<211> 18

<212> DNA

<213> Homo sapiens

<400> 490

cacaaattcc atttccca 18

<210> 491

<211> 21

<212> DNA

<213> Homo sapiens

<400> 491

tcaatagggtg atccaacatt t 21

<210> 492

<211> 18

<212> DNA

<213> Homo sapiens

<400> 492

aaagtccac aaagggtc 18

<210> 493

<211> 18
<212> DNA
<213> Homo sapiens

<400> 493
gggtaggggg atcttttt 18

<210> 494
<211> 18
<212> DNA
<213> Homo sapiens

<400> 494
tgtggaacat tcattggc 18

<210> 495
<211> 18
<212> DNA
<213> Homo sapiens

<400> 495
gtcctgggaa agatggaa 18

<210> 496
<211> 18
<212> DNA
<213> Homo sapiens

<400> 496
tcaaagcgtc tcccataa 18

<210> 497
<211> 18
<212> DNA
<213> Homo sapiens

<400> 497
tctttcgctg tacttggc 18

<210> 498
<211> 18
<212> DNA
<213> Homo sapiens

<400> 498
tgggaggtca gagtgatg 18

<210> 499
<211> 19
<212> DNA
<213> Homo sapiens

<400> 499
ggacagtgta tgtgttggg 19

<210> 500
<211> 18
<212> DNA
<213> Homo sapiens

<400> 500
aggcagctgt ttttgtga 18

<210> 501
<211> 18
<212> DNA
<213> Homo sapiens

<400> 501
cttcttgagt cccgtgtg 18

<210> 502
<211> 19
<212> DNA
<213> Homo sapiens

<400> 502
caaccgagaa tcctctagc 19

<210> 503
<211> 18
<212> DNA
<213> Homo sapiens

<400> 503
gctgggagag aatcacaa 18

<210> 504
<211> 18
<212> DNA
<213> Homo sapiens

<400> 504
gctttgcaga agagacca 18

<210> 505
<211> 18
<212> DNA
<213> Homo sapiens

<400> 505
acgctgtcag gtcacact 18

<210> 506
<211> 18
<212> DNA
<213> Homo sapiens

<400> 506

ggaggatgct caggtgat	18
<210> 507	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 507	
tagggggatc tttttcca	18
<210> 508	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 508	
gagcaatttg aaaagcca	18
<210> 509	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 509	
atggtccagc tcctctgt	18
<210> 510	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 510	
atagagcacc ccatctcc	18
<210> 511	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 511	
aacattgctg ttagccca	18
<210> 512	
<211> 18	
<212> DNA	
<213> Homo sapiens	
<400> 512	
gcaatcgaaa cagcattc	18
<210> 513	
<211> 18	
<212> DNA	
<213> Homo sapiens	

<400> 513 atgagttggc agctgaag	18
<210> 514 <211> 18 <212> DNA <213> Homo sapiens	
<400> 514 aatgaaggtc ttgcctcc	18
<210> 515 <211> 20 <212> DNA <213> Homo sapiens	
<400> 515 gaggagaaga tccacaagcg	20
<210> 516 <211> 20 <212> DNA <213> Homo sapiens	
<400> 516 tctctggggc atactgaacc	20
<210> 517 <211> 18 <212> DNA <213> Homo sapiens	
<400> 517 ctgagctttt ggcactgt	18
<210> 518 <211> 18 <212> DNA <213> Homo sapiens	
<400> 518 ctgctaggtg acagcagg	18
<210> 519 <211> 20 <212> DNA <213> Homo sapiens	
<400> 519 tgtatgagtc tggagggtgt	20
<210> 520 <211> 18 <212> DNA	

<213> Homo sapiens

<400> 520
acacctggct gaggaat 18

<210> 521
<211> 18
<212> DNA
<213> Homo sapiens

<400> 521
gcaggggacg tgataata 18

<210> 522
<211> 18
<212> DNA
<213> Homo sapiens

<400> 522
ttttgcttcc taccatgc 18

<210> 523
<211> 18
<212> DNA
<213> Homo sapiens

<400> 523
aaaattgtga gcacctcc 18

<210> 524
<211> 23
<212> DNA
<213> Homo sapiens

<400> 524
tttatattta aagtggcttt gtt 23

<210> 525
<211> 18
<212> DNA
<213> Homo sapiens

<400> 525
gtgcaaagcc cacagtat 18

<210> 526
<211> 18
<212> DNA
<213> Homo sapiens

<400> 526
aggaaaatgc aagagcag 18

<210> 527

<211> 20
<212> DNA
<213> Homo sapiens

<400> 527
ccactgaatt gcatactttg 20

<210> 528
<211> 18
<212> DNA
<213> Homo sapiens

<400> 528
tctgggtcca gtctgcta 18

<210> 529
<211> 18
<212> DNA
<213> Homo sapiens

<400> 529
agattttggg gagtcagg 18

<210> 530
<211> 17
<212> DNA
<213> Homo sapiens

<400> 530
gcgctcaagc aattctc 17

<210> 531
<211> 18
<212> DNA
<213> Homo sapiens

<400> 531
caagccccaa agtagtca 18

<210> 532
<211> 18
<212> DNA
<213> Homo sapiens

<400> 532
gaatcatcca atccacga 18

<210> 533
<211> 18
<212> DNA
<213> Homo sapiens

<400> 533
agcctccagg tgactacc 18

<210> 534
<211> 18
<212> DNA
<213> Homo sapiens

<400> 534
gaaggacatg gtcagcag 18

<210> 535
<211> 18
<212> DNA
<213> Homo sapiens

<400> 535
atgctttcag cattttcg 18

<210> 536
<211> 18
<212> DNA
<213> Homo sapiens

<400> 536
tgatccgtgg tagggtta 18

<210> 537
<211> 18
<212> DNA
<213> Homo sapiens

<400> 537
gtcggattgg tttcaca 18

<210> 538
<211> 19
<212> DNA
<213> Homo sapiens

<400> 538
ttttatggga atttcagcc 19

<210> 539
<211> 21
<212> DNA
<213> Homo sapiens

<400> 539
tttgaaaag aacagaaatg t 21

<210> 540
<211> 19
<212> DNA
<213> Homo sapiens

<400> 540

ggctagtctt tcctgaacc 19

<210> 541
<211> 18
<212> DNA
<213> Homo sapiens

<400> 541
ccttaatgcc cctgattc 18

<210> 542
<211> 18
<212> DNA
<213> Homo sapiens

<400> 542
gcgtttacaa gctgagga 18

<210> 543
<211> 18
<212> DNA
<213> Homo sapiens

<400> 543
tcaagcttgc tttctcaa 18

<210> 544
<211> 18
<212> DNA
<213> Homo sapiens

<400> 544
gtagcccagc aagtgtct 18

<210> 545
<211> 18
<212> DNA
<213> Homo sapiens

<400> 545
cctggctgga gataggat 18

<210> 546
<211> 18
<212> DNA
<213> Homo sapiens

<400> 546
cttccccctct gcctatgt 18

<210> 547
<211> 18
<212> DNA
<213> Homo sapiens

<400> 547
ggcacgtact tcctacca 18

<210> 548
<211> 18
<212> DNA
<213> Homo sapiens

<400> 548
ggtgcttctt acaggcaa 18

<210> 549
<211> 16
<212> DNA
<213> Homo sapiens

<400> 549
acccaggctg gtgtgt 16

<210> 550
<211> 23
<212> DNA
<213> Homo sapiens

<400> 550
actgagttaa ttatcactcc cct 23

<210> 551
<211> 18
<212> DNA
<213> Homo sapiens

<400> 551
gatgcatttt gcttcacc 18

<210> 552
<211> 21
<212> DNA
<213> Homo sapiens

<400> 552
tctgctttta gagctgtag c 21

<210> 553
<211> 19
<212> DNA
<213> Homo sapiens

<400> 553
tcaagcttca aagagcaga 19

<210> 554
<211> 18
<212> DNA

<213> Homo sapiens

<400> 554

ggagtacatc ccaggacc

18

<210> 555

<211> 19

<212> DNA

<213> Homo sapiens

<400> 555

tggtgctttt aaatccaga

19

<210> 556

<211> 21

<212> DNA

<213> Homo sapiens

<400> 556

ctcccttact tacttgcatt g

21

<210> 557

<211> 18

<212> DNA

<213> Homo sapiens

<400> 557

tcttctccca gggaatct

18

<210> 558

<211> 18

<212> DNA

<213> Homo sapiens

<400> 558

tttatgtccc ctgagcac

18

<210> 559

<211> 19

<212> DNA

<213> Homo sapiens

<400> 559

tccctggcta tcttgaatc

19

<210> 560

<211> 16

<212> DNA

<213> Homo sapiens

<400> 560

cttgactggg tccacg

16

<210> 561

<211> 20
<212> DNA
<213> Homo sapiens

<400> 561
cgagacgcca gtagatacca

20

<210> 562
<211> 20
<212> DNA
<213> Homo sapiens

<400> 562
catcctccat gcctttcagt

20

<210> 563
<211> 20
<212> DNA
<213> Homo sapiens

<400> 563
agttccagag aacgagacgc

20

<210> 564
<211> 20
<212> DNA
<213> Homo sapiens

<400> 564
cttgtcatcc tccatgcctt

20

<210> 565
<211> 20
<212> DNA
<213> Homo sapiens

<400> 565
gagcgtgaga ggttgaggag

20

<210> 566
<211> 20
<212> DNA
<213> Homo sapiens

<400> 566
aaacaaactc cagacgcacc

20

<210> 567
<211> 24
<212> DNA
<213> Homo sapiens

<400> 567
ctgaaccact acctgtatga cctg

24

<210> 568
<211> 25
<212> DNA
<213> Homo sapiens

<400> 568
ctaactactt actcctacag ggccc 25

<210> 569
<211> 23
<212> DNA
<213> Homo sapiens

<400> 569
gaagcatttc aatacttta ctg 23

<210> 570
<211> 20
<212> DNA
<213> Homo sapiens

<400> 570
ccactccagt gcacccaatc 20

<210> 571
<211> 20
<212> DNA
<213> Homo sapiens

<400> 571
cttctcctgg ccactctgac 20

<210> 572
<211> 21
<212> DNA
<213> Homo sapiens

<400> 572
ggtttacctt tgaatcccag c 21

<210> 573
<211> 22
<212> DNA
<213> Homo sapiens

<400> 573
tgaggatgaa tgagcacata gg 22

<210> 574
<211> 21
<212> DNA
<213> Homo sapiens

<400> 574

tttgtggtcc attgagtagg c	21
<210> 575	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 575	
aggggaagga atgtgcttgg	20
<210> 576	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 576	
ttcggctgag cgggcagtgt	20
<210> 577	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 577	
attgaaggtc ctccaaaaga atgctg	26
<210> 578	
<211> 30	
<212> DNA	
<213> Homo sapiens	
<400> 578	
agaacgtcaa catatctttt tgggggacac	30
<210> 579	
<211> 22	
<212> DNA	
<213> Homo sapiens	
<400> 579	
ttgtatttga ggactttgct cg	22
<210> 580	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 580	
cggtaccatc ctcctcttcc	20
<210> 581	
<211> 20	
<212> DNA	
<213> Homo sapiens	

<400> 581
tttttgcctc atctatgccc 20

<210> 582
<211> 20
<212> DNA
<213> Homo sapiens

<400> 582
gggtgacaga gcaagactcc 20

<210> 583
<211> 18
<212> DNA
<213> Homo sapiens

<400> 583
ttgctcaagt tctcctgg 18

<210> 584
<211> 18
<212> DNA
<213> Homo sapiens

<400> 584
accttgtttt gaggggag 18

<210> 585
<211> 18
<212> DNA
<213> Homo sapiens

<400> 585
cttggctatt tggacagc 18

<210> 586
<211> 18
<212> DNA
<213> Homo sapiens

<400> 586
gggcatttac tcacttgc 18

<210> 587
<211> 19
<212> DNA
<213> Homo sapiens

<400> 587
cttgtgtcag ttgtcaggg 19

<210> 588
<211> 19
<212> DNA

<213> Homo sapiens

<400> 588
tggaattggt gtgtcttgg 19

<210> 589
<211> 18
<212> DNA
<213> Homo sapiens

<400> 589
ccagttccac tggatggt 18

<210> 590
<211> 18
<212> DNA
<213> Homo sapiens

<400> 590
atgggctgtg tttctcaa 18

<210> 591
<211> 18
<212> DNA
<213> Homo sapiens

<400> 591
ctgcctatcc ctggactt 18

<210> 592
<211> 18
<212> DNA
<213> Homo sapiens

<400> 592
agtttgtccc tagtgccc 18

<210> 593
<211> 19
<212> DNA
<213> Homo sapiens

<400> 593
caacacgtct gacatccat 19

<210> 594
<211> 16
<212> DNA
<213> Homo sapiens

<400> 594
ggatagtgca caccca 16

<210> 595

<211> 22
<212> DNA
<213> Homo sapiens

<400> 595
tgggtggtac tattgttccc at 22

<210> 596
<211> 20
<212> DNA
<213> Homo sapiens

<400> 596
agttccagcc cccttaccag 20

<210> 597
<211> 20
<212> DNA
<213> Homo sapiens

<400> 597
ggccactatc atccctgtgt 20

<210> 598
<211> 20
<212> DNA
<213> Homo sapiens

<400> 598
tttcacatgg gaagaacacg 20

<210> 599
<211> 20
<212> DNA
<213> Homo sapiens

<400> 599
acagtgcac tagggacggg 20

<210> 600
<211> 20
<212> DNA
<213> Homo sapiens

<400> 600
tgccaggatg gagataacaa 20

<210> 601
<211> 20
<212> DNA
<213> Homo sapiens

<400> 601
cctgtggcac acatatcacc 20

<210> 602
<211> 20
<212> DNA
<213> Homo sapiens

<400> 602
acaaccaaga atggagccac 20

<210> 603
<211> 20
<212> DNA
<213> Homo sapiens

<400> 603
tgctgtgtaa caagtcccca 20

<210> 604
<211> 20
<212> DNA
<213> Homo sapiens

<400> 604
tgaacggagg acctaccaag 20

<210> 605
<211> 20
<212> DNA
<213> Homo sapiens

<400> 605
gcagggtccg actcactaag 20

<210> 606
<211> 20
<212> DNA
<213> Homo sapiens

<400> 606
gctgtgagtt ccctttacgc 20

<210> 607
<211> 20
<212> DNA
<213> Homo sapiens

<400> 607
acagtgggga caaagacagg 20

<210> 608
<211> 20
<212> DNA
<213> Homo sapiens

<400> 608

tacagggcac ctcccagtag	20
<210> 609	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 609	
tcttctgtta aggtttcccc c	21
<210> 610	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 610	
tgtctcaaac ctccctctgc	20
<210> 611	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 611	
aacatatttc ctccccagcc	20
<210> 612	
<211> 19	
<212> DNA	
<213> Homo sapiens	
<400> 612	
cagtcccagc caatgagaa	19
<210> 613	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 613	
ctcctctgca tgggagaatc	20
<210> 614	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 614	
agacctggga ccagtctgtg	20
<210> 615	
<211> 20	
<212> DNA	
<213> Homo sapiens	

<400> 615
gggagacgac gtcacaagat 20

<210> 616
<211> 20
<212> DNA
<213> Homo sapiens

<400> 616
tgatggtggg aagatggtga 20

<210> 617
<211> 20
<212> DNA
<213> Homo sapiens

<400> 617
caggcatctt ctatgtgcca 20

<210> 618
<211> 20
<212> DNA
<213> Homo sapiens

<400> 618
gggaggcaca agttctttca 20

<210> 619
<211> 20
<212> DNA
<213> Homo sapiens

<400> 619
acttcgtggc actgagtgtg 20

<210> 620
<211> 20
<212> DNA
<213> Homo sapiens

<400> 620
cctttcttac ggatgaggca 20

<210> 621
<211> 20
<212> DNA
<213> Homo sapiens

<400> 621
ggctgctgag ctctttctgat 20

<210> 622
<211> 20
<212> DNA

<213> Homo sapiens

<400> 622
tgggtctctc tgcttgactt 20

<210> 623
<211> 20
<212> DNA
<213> Homo sapiens

<400> 623
tcacctactt ccagcttccg 20

<210> 624
<211> 20
<212> DNA
<213> Homo sapiens

<400> 624
agacctggga ccagtctgtg 20

<210> 625
<211> 20
<212> DNA
<213> Homo sapiens

<400> 625
ctcctctgca tgggagaatc 20

<210> 626
<211> 20
<212> DNA
<213> Homo sapiens

<400> 626
aattcaggag acctgggacc 20

<210> 627
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a BstXI-linker adapter.

<400> 627
gttttcacca cgggg 15

<210> 628
<211> 11
<212> DNA
<213> Artificial Sequence

<220>

<223> Artificial Sequence is a BstXI-linker adapter.

<400> 628
gtggtgaaga c

11

<210> 629
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 629
ccaagttctg agaagtcc

18

<210> 630
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is a primer.

<400> 630
aatacctgaa accatac

17

<210> 631
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is an allele specific oligonucleotide.

<400> 631
agactgggggt gagacgc

17

<210> 632
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence is an allele specific oligonucleotide.

<400> 632
cagactgggt tgagacgcc

19

<210> 633
<211> 24
<212> DNA
<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 633

cccgtgtgct ccgcccacca gttc

24

<210> 634

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 634

ggctcacgga gctcatcatg gactt

25

<210> 635

<211> 502

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 635

```

cccgtgtgct ccgcccacca gttccoctgc ggcgggggtc agtgtgtgga cctgcgcctg 60
cgctgcgacg gcgaggcaga ctgtcaggac cgctcagacg aggtggactg tgacgccatc 120
tgccctgccca accagttccg gtgtgcgagc ggccagtgtg tcctcatcaa acagcagtcg 180
gactccttcc cgcactgtat cgacggctcc gacgagctca tgtgtgaaat caccaagocg 240
ccctcagacg acagcccggc ccacagcagt gccatcgggc ccgtcattgg catcatcctc 300
tctctcttcg tcatgggtgg tgtctatatt gtgtgccagc gcgtggtgtg ccagcgctat 360
gcggggggcca acggggccctt cccgcacgag tatgtcagcg ggaccccgca cgtgcccctc 420
aatttcatag ccccgggcgg ttcccagcat ggccccttca caggcatcgc atgcggaaag 480
tccatgatga gctccgtgag cc 502

```

<210> 636

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 636

agcgaggcca ccatccacag g

21

<210> 637

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 637

tcgctggtcg gcataatcaa t

21

<210> 638

<211> 501

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a primer.

<400> 638

```

agcagagcca ccattccacag gatctccctg gagactaaca acaacgatgt ggctatccca 60
ctcacgggtg tcaaagaggc ctctgcactg gactttgatg tgtccaacaa tcacatctac 120
tggactgatg ttagcctcaa gacgatcagc cgagccttca tgaatgggag ctcagtggag 180
cacgtgattg agtttggcct cgactaccct gaaggaatgg ctgtggactg gatgggcaag 240
aacctctatt gggcggacac agggaccaac aggattgagg tggcccggtt ggatgggcag 300
ttccggcagg tgcttgtgtg gagagacctt gacaacccca ggtctctggc tctggatcct 360
actaaaggct acatctactg gactgagtgg ggtggcaagc caaggattgt gcgggccttc 420
atggatggga ccaattgtat gacactggta gacaagggtg gccgggcca cgacctcacc 480
attgattatg ccgaccagcg a 501

```

<210> 639

<211> 26

<212> RNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

<400> 639

raguacagcu ucuugccaac ccaguc

26

<210> 640

<211> 26

<212> RNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

<400> 640

ruccuccagg ucgaugguca gcccau

26

<210> 641

<211> 26

<212> RNA

<213> Artificial Sequence

<220>

<223> Artificial Sequence is a Zmax1 oligonucleotide.

<400> 641
rgucugaguc cgagucaaa uccagg

26

<210> 642
<211> 3081
<212> DNA
<213> Homo sapiens

<400> 642
cccgccagcc cagcccagcc caaccctact ccctccccac gccagggcag cagccgttgc 60
tcagagagaa ggtggaggaa gaaatccaga ccctagcacg cgcgcacccat catggaccat 120
tatgattctc agcaaaccaa cgattacatg cagccagaag aggactggga ccgggacctg 180
ctcctggacc cggcctggga gaagcagcag agaaagacat tcacggcatg gtgtaactcc 240
cacctccgga aggcggggac acagatcgag aacatcgaag aggacttccg ggatggcctg 300
aagctcatgc tgctgtcgga ggtcatctca ggtgaacgct tggccaagcc agagcgaggg 360
aagatgagag tgcacaagat ctccaacgct aacaaggccc tggatttcat agccagcaaa 420
ggcgtcaaac ttgtgtccat cggagccgaa gaaatcgtgg atgggaatgt gaagatgacc 480
ctgggcatga tctggacccat catcctgcgc tttgccatcc aggacatctc cgtggaagag 540
acttcagcca aggaagggtc gctcctgtgg tgtcagagaa agacagcccc ttacaaaaat 600
gtcaacatcc agaacttcca cataagctgg aaggatggcc tcggcttctg tgctttgatc 660
caccgacacc ggcccagact gattgactac gggaagctgc ggaaggatga tccactcaca 720
aatctgaata cggcttttga cgtggcagag aagtacctgg acatccccc aa gatgctggat 780
gccgaagaca tcgttggaa cgtggcagag gatgagaaag ccatcatgac ttacgtgtct 840
agcttctacc acgccttctc tggagcccag aaggcggaga cagcagccaa tcgcatctgc 900
aaggtgttgg ccgtcaacca ggagaacgag cagcttatgg aagactacga gaagctggcc 960
agtgatctgt tggagtggat ccgccgcaca atcccgtggc tggagaaccg ggtgcccag 1020
aacaccatgc atgccatgca acagaagctg gaggacttcc gggactaccg gcgcctgcac 1080
aagccgcca aggtgcagga gaagtgccag ctggagatca acttcaacac gctgcagacc 1140
aagctgcggc tcagcaaccg gcctgccttc atgccctctg agggcaggat ggtctcggac 1200
atcaacaatg cctggggctg cctggagcag gtggagaagg gctatgagga gtggttgctg 1260
aatgagatcc ggaggctgga ggcactggac cacctggcag agaagttccg gcagaaggcc 1320
tccatccacg aggcctggac tgacggcaaa gaggccatgc tgcgacagaa ggactatgag 1380
accgccaccc tctcggagat caaggccttg ctcaagaagc atgaggcctt cgagagtgcac 1440
ctggctgccc accaggaccg tgtggagcag attgccgcca tcgcacagga gctcaatgag 1500
ctggactatt atgactcacc cagtgtcaac gcccgctggc aaaagatctg tgaccagtgg 1560
gacaatctgg gggccctaac tcagaagcga agggaagctc tggagcggac cgagaaactg 1620
ctggagacca ttgaccagct gtacttggag tatgccaaag gggctgcacc cttcaacaac 1680
tggatggagg gggccatgga ggacctgcag gacaccttca ttgtgcacac cattgaggag 1740
atccagggac tgaccacagc ccatgagcag ttcaaggcca ccctccctga tgccgacaag 1800
gagcgcctgg ccatcctggg catccacaat gaggtgtcca agattgtcca gacctaccac 1860
gtcaatatgg cgggcaccaa cccctacaca accatcacgc ctcaggagat caatggcaaa 1920
tgggaccacg tgcggcagct ggtgcctcgg agggaccaag ctctgacgga ggagcatgcc 1980
cgacagcagc acaatgagag gctacgcaag cagtttggag cccaggccaa tgtcatcggg 2040
ccctggatcc agaccaagat ggaggagatc gggaggatct ccattgagat gcatgggacc 2100
ctggaggacc agctcagcca cctgcggcag tatgagaaga gcatcgtcaa ctacaagcca 2160
aagattgatc agctggaggg cgaccaccag ctcatccagg aggcgctcat cttcgacaac 2220
aagcacacca actacaccat ggagcacatc cgtgtgggct gggagcagct gctcaccacc 2280
atcgccagga ccatcaatga ggtagagaac cagatcctga cccgggatgc caagggcac 2340
agccaggagc agatgaatga gttccgggcc tccttcaacc actttgaccg ggatcactcc 2400
ggcacactgg gtcccagga gttcaaagcc tgcctcatca gcttgggtta tgatattggc 2460
aacgaccccc agggagaagc agaatttgcc cgcacatgga gcattgtgga cccaaccgc 2520
ctgggggtag tgacattcca ggccttcatt gacttcatgt cccgcgagac agccgacaca 2580
gatacagcag accaagtcac ggcttccttc aagatcctgg ctggggacaa gaactacatt 2640
accatggacg agctgcgccg cgagctgcca cccgaccagg ctgagtactg catcgcgcgg 2700

atggccccct	acaccggccc	cgactccgtg	ccaggtgctc	tggactacat	gtcctttctcc	2760
acggcgctgt	acggcgagag	tgacctctaa	tccaccccg	ccggccgccc	tcgtcttggtg	2820
cgccgtgccc	acagatgtga	aatgaatgta	atctaataga	agcctaataca	gcccaccatg	2880
ttctccactg	aaaaatcctc	tttctttggg	gtttttcttt	ctttcttttt	tgattttgca	2940
ctggacgggtg	acgtcagcct	gtacaggctc	ccaggggtgg	cgtcaaatgc	tattgaaatt	3000
gcgctgaatc	gtatgctttt	tccttttgat	aaataaacia	tgtaaaaatg	tttcaaaaac	3060
ctaataaaat	aaataaatac	g				3081

<210> 643

<211> 1324

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(1324)

<223> n = A,T,C or G

<400> 643

ggccgcccgg	cgccccagc	agnccgagcc	ggggcgcaca	gncggggcgc	agcccgcgcc	60
ccccgcccgg	attgacatga	tgtttccaca	aagcaggcat	tcgggctcct	cgcacctacc	120
ccagcaactc	aaattcacca	cctcggactc	ctgcgaccgc	atcaaagacg	aatttcagct	180
actgcaagct	cagtaccaca	gcctcaagct	cgaatgtgac	aagttggcca	gtgagaagtc	240
agagatgcag	cgtcactatg	tgatgtacta	cgagatgtcc	tacggcttga	acatcgagat	300
gcacaaacag	gctgagatcg	tcaaaaggct	gaacgggatt	tgtgcccagg	tcctgccccta	360
cctctcccaa	gagcaccagc	agcaggctct	gggagccatt	gagagggcca	agcagggtcac	420
cgctcccag	ctgaactcta	tcatccgaca	gcagctccaa	gcccaccagc	tgtcccagct	480
gcaggccctg	gccctgccct	tgaccccact	accctgtggg	ctgcagccgc	cttcgctgcc	540
ggcggtcagc	gcaggcaccg	gcctcctctc	gctgtccgcg	ctgggttccc	aggcccacct	600
ctccaaggaa	gacaagaacg	ggcacgatgg	tgacaccac	caggaggatg	atggcgagaa	660
gtcggattag	cagggggccg	ggacagggag	gttgggaggg	gggacagagg	ggagacagag	720
gcacggagag	aaaggaatgt	ttagcacaaag	acacagcgga	gctcgggatt	ggctaattctc	780
ccatagtatt	tatgggtggcg	ccggcggggc	cccagcccag	cttgacaggcc	acctctagct	840
ttcttctctac	cccattccgg	cttccctcct	cctcccctgc	agcctgggta	gggtggatacc	900
tgccctgaca	tgtgaggcaa	gctaaggcct	ggagggtcag	atgggagacc	aggtcccaag	960
ggagcaagac	ctgcgaagcg	cagcagcccc	ggcccttccc	ccgttttgaa	catgtgtaac	1020
cgacagtctg	ccctggggcca	cagccctctc	accctgggtac	tgcatgcacg	caatgctagc	1080
tgcccctttc	ccgtcctggg	caccccgagt	ctcccccgac	cccgggtccc	aggatatgctc	1140
ccacctccac	ctgccccact	caccacctct	gctagtcca	gacacctcca	cgcccacctg	1200
gtcctctccc	atcgcccaca	aaaggggggg	cacgagggag	gagcttagct	gagctgggag	1260
gagcaggggtg	aggggtgggcg	accagggatt	ccccctcccc	ttcccaaata	aagatgaggg	1320
tact						1324

<210> 644

<211> 2377

<212> DNA

<213> Homo sapiens

<400> 644

ggtgacaaaag	agccaacaga	gacaatagga	gacttgtcaa	tttgtcttga	tgggctacag	60
ttagagtctg	aagttgttac	caatgggtgaa	actacatggt	cagaaagtgc	ttctcagaat	120
gatgatggct	ccagatccaa	ggatgaaaca	agagtgcagc	caaattggatc	agatgacctt	180
gaagatgcag	gagctgggtg	aaataggaga	gtcagtggga	ataattctcc	atcactctca	240
aatgggtggtt	ttaaaccctc	tagacctcca	agaccttcac	gaccaccacc	acccacccca	300

cgtagaccag	catctgtcaa	tggttcacca	tctgccactt	ctgaaagtga	tgggtctagt	360
acaggctctc	tgccgccgac	aaatacaaat	acaaatacat	ctgaaggagc	aacatctgga	420
ttaataattc	ctcttactat	atctggaggc	tcaggcccta	ggccattaaa	tcctgtaact	480
caagctccct	tgccacctgg	ttggggagcag	agagtggacc	agcacgggcg	agtttactat	540
gtagatcatg	ttgagaaaaa	aacaacatgg	gatagaccag	aacctctacc	tcctggctgg	600
gaacggcggg	ttgacaacat	gggacgtatt	tattatgttg	accatttcac	aagaacaaca	660
acgtggcaga	ggccaacact	ggaatccgtc	cggaaactatg	aacaatggca	gctacagcgt	720
agtcagcttc	aaggagcaat	gcagcagttt	aaccagagat	tcatttatgg	gaatcaagat	780
ttatttgcta	catcacaaag	taaagaattt	gatcctcttg	gtccattgcc	acctggatgg	840
gagaagagaa	cagacagcaa	tggcagagta	tatttcgtca	accacaacac	acgaattaca	900
caatgggaag	accccagaag	tcaagggtcaa	ttaaatgaaa	agcccttacc	tgaagggttg	960
gaaatgagat	tcacagtggg	tgggaattcca	tattttgtgg	accacaatag	aagaactacc	1020
acctatatag	atccccgcac	aggaaaatct	gccctagaca	atggacctca	gatagcctat	1080
gttcgggact	tcaaagcaaa	ggttcagtat	ttccggttct	ggtgtcagca	actggccatg	1140
ccacagcaca	taaagattac	agtgacaaga	aaaacattgt	ttgaggattc	ctttcaacag	1200
ataatgagct	tcagtcccca	agatctgcga	agacgtttgt	gggtgatttt	tccaggagaa	1260
gaagggttag	attatggagg	tgtagcaaga	gaatggttct	ttcttttgtc	acatgaagtg	1320
ttgaacccaa	tgtattgcct	gtttgaatat	gcagggaagg	ataactactg	cttgacagata	1380
aaccccgctt	cttacatcaa	tccagatcac	ctgaaatatt	ttcgttttat	tggcagattt	1440
attgccatgg	ctctgttcca	tgggaaattc	atagacacgg	gtttttcttt	accattctat	1500
aagcgtatct	tgaacaaaacc	agttggactc	aaggatttag	aatctattga	tccagaattt	1560
tacaattctc	tcatctgggt	taaggaaaac	aatattgagg	aatgtgattt	ggaaatgtac	1620
ttctccgttg	acaaagaaat	tctaggtgaa	attaagagtc	atgatctgaa	acctaattgt	1680
ggcaatatct	ttgtaacaga	agaaaataaa	gaggaatata	tcagaatggg	agctgagtgg	1740
agggtgtctc	gagggtgttg	agaacagaca	caagctttct	ttgaaggctt	taatgaaatt	1800
cttccccagc	aatatattgca	atactttgat	gcaaaggaat	tagaggctct	tttatgtgga	1860
atgcaagaga	ttgatttgaa	tgactggcaa	agacatgcca	tctaccgtca	ttatgcaagg	1920
accagcaaac	aatcatgtg	gttttggcag	tttgttaaag	aaattgataa	tgagaagaga	1980
atgagacttc	tgcagtttgt	tactggaacc	tgccgattgc	cagtaggagg	atttgctgat	2040
ctcatgggga	gcaatggacc	acagaaattc	tgcaattgaa	aagttgggaa	agaaaattgg	2100
ctaccagaaa	gtcatacctg	ttttaatcgc	ctggacctgc	caccatacaa	gagctatgag	2160
caactgaagg	aaaagctgtt	gtttgccata	gaagaaacag	aaggatttgg	acaagagtaa	2220
cttctgagaa	cttgcaaccat	gaatgggcaa	gaacttattt	gcaatgtttg	tccttctctg	2280
cctgttgcac	atcttgtaaa	attggacaat	ggctctttag	agagttatct	gagtgtaaagt	2340
aaattaatgt	tctcatttaa	aaaaaaaaaa	aaaaaaa			2377

<210> 645

<211> 1295

<212> DNA

<213> Homo sapiens

<400> 645

gggccgccgc	ccaccggggc	cttgccctcta	cctcagtcgt	tgccccccga	ttttcggtctg	60
gagccacagg	ccccggccct	cagccccgc	tctagcttcg	ccagtagctc	ggccagcgac	120
gcgagcaagc	cgtccagccc	ccggggcagc	ctgctgctgg	acggggcggg	ggctggcgga	180
gctggaggta	gccggccctg	cagcaatcgc	accagcggca	tcagcatggg	ctacgaccag	240
cgccacggga	gcccccttgc	agcggggccg	tgccctgttg	gccccccct	ggccggagca	300
ccggcaggct	attctcccgg	aggggtcccg	tccgcctacc	cggagctcca	cgccgcccctg	360
gaccgattgt	acgctcagcg	gcccgcgggg	ttcggtgcc	aggaaagccg	ccactcgtat	420
cccccgggcc	tgggcagccc	tggagctcta	gccggggccc	gagtgggagc	ggcggggccc	480
ttggagagac	ggggggcgca	accgggacga	cactctgtga	ccggctacgg	ggactgcgcc	540
gtggggcgccc	ggtaccagga	cgagctaaca	gctttgcttc	gcctgacggg	gggcaccggg	600
gggcgagaa	ccggagcccc	cggagaaccc	tgggggattg	agccgtcggg	tctggaggag	660
ccaccaggtc	ctttcgttcc	ggaggccgcc	cgggcccggg	tgcgggagcc	agaggccagg	720

gaggactact	tcggcacctg	tatcaagtgc	aacaaaggca	tctatgggca	gagcaatgcc	780
tgccaggccc	tggacagcct	ctaccacacc	cagtgtcttg	tttgtctgctc	ttgtgggcca	840
acttttgcgtt	gcaaggcttt	ctacagtgtc	aatggctctg	tgtactgtga	ggaagattat	900
ctgttttcag	ggtttcagga	ggcagctgag	aaatgctgtg	tctgtgggtca	cttgattttg	960
gagaagatcc	tacaagcaat	ggggaagtcc	tatcatccag	gctgtttccg	atgcattgtt	1020
tgcaacaagt	gcttggatgg	catccccctc	acagtggact	tctccaacca	agtatactgt	1080
gtcaccgact	accacaaaaa	ttatgtctct	aagtgtgcag	cctgtggcca	acccatcctc	1140
ccctctgagg	gctgtgagga	catcgtgagg	gtgatatcca	tggaccggga	ttatcacttt	1200
gagtgtacc	actgtgagga	ctgccggatg	cagctgagtg	atgaggaagg	ctgctgctgt	1260
ttccctctgg	atgggcactt	gctctgccat	ggttg			1295

<210> 646

<211> 3411

<212> DNA

<213> Homo sapiens

<400> 646

ggggccgggg	tcccgccacc	accgcgcgcg	ggacagattg	attcactttg	gagctgtaag	60
tactgatgta	ttaggggtgca	gcgctcattg	ttcattgacg	cagagtccca	aaatgaatat	120
ccaagagcag	ggtttccctt	tggacctcgg	agcaagtttc	accgaagatg	ctccccgacc	180
cccagtgcct	ggtgaggagg	gagaactggg	gtccacagac	ccgaggcccg	ccagctacag	240
tttctgctcc	gggaaagggtg	ttggcattaa	aggtgagact	tcgacggcca	ctccgaggcg	300
ctcgatctg	gacctggggg	atgagcctga	gggcagtgcc	tccccacccc	caccatactt	360
gaagtgggct	gagtcactgc	attccctgct	ggatgaccaa	gatgggataa	gcctgttcag	420
gactttcctg	aagcaggagg	gctgtgccga	cttgcctggac	ttctggtttg	cctgcactgg	480
cttcaggaag	ctggagccct	gtgactcgaa	cgaggagaag	aggctgaagc	tggcgagagc	540
catctaccga	aagtacattc	ttgataacaa	tggcatcgtg	tcccggcaga	ccaagccagc	600
caccaagagc	ttcataaagg	gctgcatcat	gaagcagctg	atcgatcctg	ccatgtttga	660
ccaggccccag	accgaaatcc	aggccactat	ggaggaaaac	acctatccct	ccttccttaa	720
gtctgatatt	tatttggaa	atagcaggac	aggctcggag	agccccaaag	tctgtagtga	780
ccagagctct	gggtcaggga	cagggaaggg	catatctgga	tacctgccga	ccttaaataa	840
agatgaggaa	tggaaagtgtg	accaggacat	ggacgaggac	gatggcagag	acgctgctcc	900
ccccggaaga	ctccctcaga	agctgctcct	ggagacagct	gccccgaggg	tctcctccag	960
tagacgggtac	agcgaaggca	gagagttcag	gtatggatcc	tggcggggagc	cagtcaaccc	1020
ctattatgtc	aatgccggct	atgccctggc	cccagccacc	agtgccaacg	acagcgagca	1080
gcagagcctg	tccagcgatg	cagacaccct	gtccctcacg	gacagcagcg	tggatgggat	1140
ccccccatac	aggatccgta	agcagcaccg	caggagagatg	caggagagcg	cgcagggtcaa	1200
tgggcgggtg	cccctacctc	acattccccg	cacgtaccgg	gtgccgaagg	aggtccgcgt	1260
ggagcctcag	aagtctgcgg	aggagctcat	ccaccgcctg	gaggctgtgc	agcgcacgcg	1320
ggaggccgag	gagaagctgg	aggagcggct	gaagcgcgtg	cgcattggagg	aggaagggtga	1380
ggacggcgat	ccatcgtcag	ggcccccagg	gccgtgtcac	aagctgcctc	ccgccccgc	1440
ttggcaccac	ttcccgcctc	gcttgtgttg	gacatgggct	tgtgccgggc	tccgggatgc	1500
acacgaggag	aaccctgaga	gcacctctga	cgagcacgta	cagcgtgtgc	tgaggacaac	1560
tggccgcccag	tcgcctgggc	ctggccatcg	ctccccggac	agtgggcacg	tggccaagat	1620
gccagtggca	ctgggggggtg	ccgcctcggg	gcacgggaag	cacgtacca	agtcaggggc	1680
gaagctggac	gcggccggcc	tgcaccacca	ccgacacgtc	caccaccacg	tccaccacag	1740
cacagcccg	cccaaggagc	agggtggaggc	cgaggccacc	cgcaggggcc	agagcagctt	1800
cgcttggggc	ctggaaccac	acagccatgg	ggcaaggtcc	cgaggctact	cagagagtgt	1860
tggcgctgcc	cccaacgcca	gtgatggcct	cgcccacagt	gggaagggtg	gcgttgctgtg	1920
caaaagaaat	gccaaagaag	ctgagtcggg	gaagagcgcc	agcaccgagg	tgccagggtgc	1980
ctcgagggat	gcggagaaga	accagaaaaat	catgcagtgg	atcattgagg	gggaaaagga	2040
gatcagcagg	caccgcagga	ccggccacgg	gtcttcgggg	acgaggaagc	cacagcccca	2100
tgagaactcc	agaccyttgt	cccttgagca	cccttggggc	ggccctcagc	tccggacctc	2160
cgtgcagccc	tcccacctct	tcacccaaga	ccccaccatg	ccaccaccac	cagctcccaa	2220

ccccctaacc	cagctggagg	aggcgcgcgcg	acgtctggag	gaggaagaaa	agagagccag	2280
ccgagcacc	tccaagcaga	ggtatgtgca	ggaggttatg	cggcggggac	gcgctgcgt	2340
caggccagcg	tgcgcgccgg	tgctgcacgt	ggtaccagcc	gtgtcggaca	tggagctctc	2400
cgagacagag	acaagatcgc	agaggaaggt	ggcgggcg	agtgccagc	cgtgtgacag	2460
catcgttgtg	gcgtactact	tctgcgggga	acccatcccc	taccgcaccc	tggtagggg	2520
ccgcgctgtc	accctgggccc	agttcaagga	gctgctgacc	aaaaaggcca	gctacagata	2580
ctacttcaag	aaagtgcgcg	acgagtttga	ctgtggggtg	gtgtttgagg	aggttcgaga	2640
ggacgaggcc	gtcctgccc	tctttgagga	gaagatcatc	ggcaaagtgg	agaagggtga	2700
ctgataggct	ggtgggctgg	ccgctgtgcc	aggcgaggcc	cttggcgggc	acgggtgtca	2760
cggccaggca	gatgacctcg	tactcaggag	cccgatgggg	aacagtgttg	ggtgtaccac	2820
ccatccctgt	ggtctacccg	tgtctagagg	caggtagggg	gtccctccaa	gtggtccaca	2880
agcttctgtc	ctgcccccaa	ggaggcagcc	tggaccactc	ctcatagcaa	tacttggagg	2940
gcccagccca	agtgaggcag	ccgaggtccc	tgctgccagc	ttcagggtgac	cccccccat	3000
cccccgccac	ctcccttggg	cacgtgtgct	gggatctact	ttccctctgg	gatttgccca	3060
cgtacccagg	tctggctggg	gcccaggccc	ggatgcagag	gcctgcaggg	cctctgtcaa	3120
ttgtacgcgc	caccaagtgc	cttcaacaca	gcttgtctct	tgccctgccac	tgtgtgaatc	3180
ggcgacggag	cactgcacct	gcctccagcc	gccggctgtg	cagtcctggg	tcctcctttc	3240
tgagggcccg	tgtaaatatg	tacatttctc	aggctagggc	cagcaggggc	tgcccagatc	3300
tgtttttcat	gcgatgacac	ttgtacaatt	atcttttcaa	aggtaacttg	ataataatga	3360
aataaaaactg	tttttgaacc	tgaataaaaa	aaaaaaaaaa	aaaaaaaaaa	a	3411

<210> 647

<211> 3140

<212> DNA

<213> Homo sapiens

<400> 647

ggctgcgagt	acctccatgg	tcccgggtggc	tgtgacggcg	gcagtggcgc	ctgtcctgtc	60
cataaacagc	gatttctcag	atttgcggga	aattaaaaag	caactgctgc	ttattgcggg	120
ccttaccg	gagcggggcc	tactacacag	tagcaaatgg	tcggcgagg	tggctttctc	180
tctccctgca	ttgcctctgg	ccgagctgca	accgcctccg	cctattacag	aggaagatgc	240
ccaggatatg	gatgcctata	ccctggccaa	ggcctacttt	gacgttaaag	agtatgatcg	300
ggcagcacat	ttcctgcatg	gctgcaatag	caagaaagcc	tattttctgt	atatgtattc	360
cagatatctg	tctggagaaa	aaaagaagga	cgatgaaaca	gttgatagct	taggccccct	420
ggaaaaagga	caagtgaaaa	atgaggcgct	tagagaattg	agagtggagc	tcagcaaaaa	480
acaccaagct	cgagaacttg	atggatttgg	actttatctg	tatggtgtgg	tgcttcgaaa	540
actggacttg	gttaaagagg	ccattgatgt	gtttgtggaa	gctactcatg	ttttgccctt	600
gcattgggga	gcctggttag	aactctgtaa	cctgatcaca	gacaaaagaga	tgctgaagtt	660
cctgtctttg	ccagacacct	ggatgaaaga	gttttttctg	gctcatatat	acacagagtt	720
gcagttgata	gaggaggccc	tgcaaaaagta	tcagaatctc	attgatgtgg	gcttctctaa	780
gagctcgtat	attgtttccc	aaattgcagt	tgccatcac	aatatcagag	atattgacaa	840
agccctctcc	atttttaatg	agctaaggaa	acaagaccct	tacaggattg	aaaatatgga	900
cacattctcc	aaccttcttt	atgtcaggag	catgaaatcg	gagttgagtt	atctggctca	960
taacctctgt	gagattgata	aataccgtgt	agaaacgtgc	tgtgtaattg	gcaattatta	1020
cagtttacgt	tctcagcatg	agaaagcagc	cttatatttc	cagagagccc	tgaaattaaa	1080
tcctcggtat	cttgggtgct	ggacactaat	gggacatgag	tacatggaga	tgaagaacac	1140
gtctgctgct	atccaggctt	atagacatgc	cattgaggtc	aacaaacggg	actacagagc	1200
ttggtatggc	ctcgggcaga	cctatgaaat	ccttaagatg	ccattttact	gcctttatta	1260
ttatagacgg	gccaccagc	ttcgacccaa	tgattctcgc	atgctgggtg	ctttaggaga	1320
atgttacgag	aaactcaatc	aactagtggg	agccaaaaag	tggtattgga	gagcttacgc	1380
cgtgggagat	gtggagaaaa	tggctctggg	gaaactggca	aagcttcatg	aacagttgac	1440
tgagtcagaa	caggctgccc	agtgttacat	caaatatata	caagatatct	attcctgtgg	1500
ggaaatagta	gaacacttgg	aggaaagcac	tgcttttcgc	tatctggccc	agtactatct	1560
taagtgc aaa	ctgtgggatg	aagcttcaac	ttgtgcacaa	aagtgttgtg	catttaatga	1620

tacccgggaa	gaaggtaagg	ccttactccg	gcaaataccta	cagcttcgga	accaaggcga	1680
gactcctacc	accgaggtgc	ctgctccctt	tttcctacct	gcttcactct	ctgctaacaa	1740
tacccccaca	cgcagagttt	ctccactcaa	cttgtcttct	gtcacgccat	agttggctac	1800
tctcaagcca	gcacattggt	agacccatct	taattaagcc	ttacctccat	gtaaagaaca	1860
gcacgtctgt	tccaaggacc	tcagctcttc	ttgtttctac	agatggcaac	agctccatag	1920
ggacagcttg	tataattacc	ttcagaggcc	aactgacaga	atcctggcag	gaacagacat	1980
tatcttgcca	gttagaagta	cttctgtctc	acttatgtcc	aaagagtggc	tatagatctt	2040
ggccttcttc	cctgaatgct	tttttttttt	ggcccccaag	aaagtccctt	ttatagcact	2100
ttagcacagg	caatgctaca	ggaacaaagt	ttcaatgctg	ctgagagtga	aagaaaggag	2160
gaaagtctgc	cactctaccc	tgagctggca	gtagggcact	gagtaccctt	aggaagaagt	2220
cagagcaatg	gatacaaatg	accttgctct	tggatttgct	gagcatgac	cctattctga	2280
tgctcagagat	taggtttaaa	tggaaatagag	ctatccattt	gttcttactc	tctagggaga	2340
caatcttcca	aaacagtttt	gggggggtct	tctaaagctt	tcaaattgga	agtaacttta	2400
ttcaactaga	gttgaataaa	agaagggcaa	aaataatctc	acagagcttg	gaactgctga	2460
tagcccttac	tgaggggcaa	agatggctat	attgttagct	atactcctac	caaagcaagc	2520
aaggagatag	gattatagat	aatttcacgg	acatttgga	ataacattgg	tgattataca	2580
gacaagaata	aactcacttc	aagctggctc	gttttaataa	atcttcaacg	taattgtcta	2640
tttttttccc	tcccatctgc	aacagaatac	atctttttca	gcctttatct	agatgaggta	2700
aagggaatca	ttcttatggg	gctcttggag	agtttcaggc	ctgtgcatgt	gtgtacagca	2760
ggaggtaata	tgctataatg	tctgctgtaa	tatatgtgca	cagtagatgc	tatggatcat	2820
tctgagctca	gggtccagac	tttattctta	ttcccagaat	tttgtgttac	gtttttacct	2880
cctaacatat	gacacttcat	cttatattaa	ggaaggttta	gaatatctaa	tacgaactga	2940
attcatttgt	tactaagcct	tctcaggcaa	gctgtatact	agttactggg	ctccactgcc	3000
atgccttttc	aagggtccca	tgggtccagaa	tgatgtttga	ttcttaattt	ttctgtccct	3060
tttataaatt	gttttaatat	ttttgctaca	tttggaattc	aataaaaaat	gtgaacaata	3120
ataaaaaaaa	aaaaaaaaaa					3140

<210> 648

<211> 3513

<212> DNA

<213> Homo sapiens

<400> 648

ccgtgtacca	ggtgctgcta	gtgggaagca	cgtgctgaa	ggaagtgcct	tccgggctgc	60
agctggagca	gttgcttct	cagagcctgc	tgacccacat	cccaacggcg	gggctgcca	120
cttcgctagg	aggaggcctg	ccttactgcc	accaggcctg	gctggatttc	cgaaggcggc	180
tggaaagctct	actacagaac	tgccaggcag	cttgtgccct	gctccagggg	gccatcgaaa	240
gtgtgaaggc	tgtgccccag	cccatggagc	ctggggagggt	cggtcagctg	ctacagcaga	300
cagaggctct	gatgcagcag	gtgctagact	cgccatggct	ggcatggcta	caatgccagg	360
ggggccggga	gctgacatgg	ctgaagcaag	aggctccaga	ggtgacctg	agcccagact	420
acaggacggc	aatggacaag	gctgacgagc	tatatgaccg	ggtggatgga	ttgctgcacc	480
aactgaccct	gcagagcaac	cagcgaatac	aggccctaga	gttgggtcca	acactggagg	540
cccgggaaaag	cggactgcac	cagattgaag	tgtggctgca	gcagggtggc	tggccagcac	600
tggaggaggc	tggggagccc	tcgctggaca	tgtgctcca	ggcccaaggc	tcttttcagg	660
agctgtacca	ggttgcccag	gagcaggctc	ggcaagggga	gaagtttctg	cagccgctga	720
ctggctggga	ggcggctgaa	ctggaccccc	ctggggcacg	ctttctggcc	ctgcgagccc	780
agctgactga	attctctagg	gctttggccc	agcgggtgcca	gcggctggcg	gatgctgaga	840
ggctgtttca	gctcttcagg	gaggccttga	cgtgggctga	ggaggggcag	cgagtgttgg	900
cagagctgga	gcaggaaacgc	ccgggggttg	tgttgacgca	gctgcagctg	cactggacca	960
ggcaccctga	cttgccctct	gcccacttcc	gaaagatgtg	ggctctggcc	acggggctgg	1020
gctcagaggc	catccgccag	gagtgccgct	gggcctgggc	gcgggtgccag	gacacctggc	1080
tggccctgga	ccaaaagctt	gaggcttcac	tgaagctacc	accggtgggc	agcacagcta	1140
gcctgtgtgt	cagccaggct	cccgtgcac	ctgccacccc	tcccctgagg	aaggcctaca	1200
gcttcgatcg	gaatctgggg	cagagtctca	gtgaacctgc	ctgccactgc	caccatgcgg	1260

ccactattgc	tgcttgccgc	agaccagagg	ctggaggagg	tgccctgccc	caggcatccc	1320
ctactgtgcc	tccaccaggc	agctctgacc	ccaggagcct	caacaggcta	cagctggtgc	1380
tggcagagat	ggtggccacg	gagcgggagt	atgtccgggc	tctagagtac	actatggaga	1440
actatttccc	cgagctggat	cgccccgatg	tgccccaggg	cctccgcggt	cagcgtgccc	1500
acctcttttg	caacctggag	aagctgcggg	acttccactg	ccacttcttc	ctgctgagc	1560
tggaggcttg	caccggcac	ccaccacgag	tggcctatgc	cttcctgcmc	cataggggtgc	1620
agtttggtg	gtacgcgctc	tacagcaaga	ataagcctcg	ctccgatgcc	ctgatgtcaa	1680
gctatgggca	caccttcttc	aaggacaagc	agcaagcact	gggggaccac	ctggacctgg	1740
cctcctacct	gctaaagccc	atccagcgca	tgggcaagta	cgcactgctg	ctgcaggagc	1800
tggcacgggc	ctgcgggggc	cccacgcagg	agctcagtgc	gctgcgggag	gcccagagcc	1860
ttgtgcactt	ccagctgcgg	cacggaaacg	acctgctggc	catggacgcc	atccagggct	1920
gtgatgttaa	cctcaaggaa	caggggcagc	tgggtgcgaca	ggatgagttt	gtggtgcgca	1980
ctgggcgcca	caagtcctgt	cgccgcctct	tcctttttga	ggagctgctg	ctcttcagca	2040
agcctcgcca	tgggcccaca	ggggttgaca	catttgccta	caagcgctcc	ttcaagatgg	2100
cagaccttgg	tctcactgag	tgctgtggga	acagcaacct	gcgcttcgag	atctggttcc	2160
gccgcgcaa	ggccagggac	acctttgtgc	tgcaggcctc	cagcctggct	atcaagcagg	2220
cctggacagc	tgacatctcc	cactgtcttt	ggaggcaggc	cgtccacaac	aaggagggtgc	2280
gcatggctga	gatgtgttcc	atgggtgtgg	ggaacaaggc	cttccgagac	attgctccca	2340
gcgaggaagc	catcaacgac	cgcaccgtca	actatgtcct	gaagtgcgca	gaagttcgct	2400
ctcgggcgtc	cattgcccga	gccccgtttg	accatgacag	cctctacctg	ggggcctcga	2460
actcccttcc	tggagacctt	gcctcttctg	ctgttctggg	gtccctcaac	ctgcacctgt	2520
acagagaccc	agctcttctg	ggtctccgct	gtcccttgta	tcccagcttc	ctagaggaag	2580
cagcactgga	ggctgaggca	gagctgggag	gccagccctc	tttgactgct	gaggactcag	2640
agatctcgtc	ccaatgcccc	tcagccagtg	gctccagtg	ctctgacagc	agctgtgtgt	2700
cagggcaggc	cctgggtagg	ggcctggagg	acttaccctg	tgtctgagcc	cgggactgga	2760
cgagcagtag	atccagcagc	ctgcagctcc	aagggaacatt	gcctctctgg	atctgctgtg	2820
accagggtgt	ggctgacacc	tgggctacct	ccaacctaca	tgtgcaacgc	tgttgactac	2880
cctttctgat	gtgtgtggcc	attggactaa	ctggcacggg	gcctctctag	ggaagtctgg	2940
ttgtagagcc	tgaataggct	cctggcccca	tgacccttcc	tcctgtcccc	agctcccatc	3000
ccagttgtgg	gttaagaata	ggctagagca	gacattgggt	gtttccatgc	tgtaggctgg	3060
tgggggacca	tgtgcctcta	ggcagtgact	aggggtgccc	caccctcag	gaagaacaca	3120
ggtgggctcc	tagcagctga	tccccaatgc	ctggccttaa	agccgagctc	agttaccata	3180
gggacaggtc	cacctctact	gggcccctcat	gcttgccttt	cctggccccc	aggcccagcc	3240
cctttttact	ggggcagttt	cgttattttg	acttgatgcc	ttttgaataa	ctttcaatag	3300
aattgtctaa	aattatctta	ctggttgtta	ggcctttggt	gtctcagaga	aggagcttag	3360
gtctttgatg	tgtgatttaa	tccttttatt	gtttataata	aaaaatagac	tgatttgtaa	3420
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3480
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaa			3513

<210> 649

<211> 3597

<212> DNA

<213> Homo sapiens

<400> 649

catgccagtt	acttcctcag	gaaaatattt	tcttgccttc	ttctttcagt	atggttttta	60
atttgggaac	agtggataac	ccaagtgtcc	cacaggccaa	ggtatattcc	aatggcagca	120
tgatocctgc	acccaaagcc	agcccetaaa	gcctaccctt	tgtgcacccg	cagcctggta	180
agtgagcttg	gctgcttgtg	aggagctaca	agtgaagag	aagttatttt	aaataaatcc	240
caaagtttga	ggcagactgt	ccaggactgt	tcccaggaag	aagcaggagt	taccacaggg	300
aaaagtctct	gacctggtcc	cctcaggccc	agctacctgc	gcccaccagc	agtgaagggt	360
gatgtactgg	cccagcatct	ccacctcccc	catgcaacca	ggtccctggg	accgtgtctc	420
ccgttgcatg	tctggcttct	gcctgtgtct	ctcctgccac	gagcatcctc	cctgtccctc	480
ctcattccac	cgtgtctctc	ctgcacacat	agcctctgtc	ccagggcgat	ttatccactt	540

gagtacagga	gctgctcaga	cctctcagcc	cagccctctg	tgactgcccc	agccccatcc	600
tacccccacc	aaagctgcct	tcctggctgt	aggagctccc	tcgtctagcc	aaggccctat	660
gggtcccat	ccgaggatcc	acaagcaatg	acttcccaaa	tgacctccac	tgcaagaaga	720
atccttacca	ctgtttccag	agccgtgaac	gatgctgtga	tggcccaggt	ctcagcacca	780
ccctctgtga	cctaaaaaga	aaagctcaat	ttccatctgt	cttctttccc	aggaccaagg	840
ggacacagta	atgtgaagtc	aaatacttaa	ccgagcaaag	ggccagtgtt	gttatcagtc	900
aaggacaaac	ctcccacctc	acagacagcc	aagcagttag	ggaaagacag	acagacatag	960
gtaggaaggt	gctctgcagg	cacaaggccc	agagaagccc	ctctccggga	acttcccctg	1020
ctccttccag	gaacagttag	cccagttagc	agtcccagcc	agctcttcaa	ggccttcaag	1080
gggtctttcc	atgactgagt	cacctccagg	agctcacctg	acccccagag	aagacctacc	1140
ccaggcagct	ccgtgccctg	gcttctcccc	atgccccaaa	tcccccccag	ccatccctcc	1200
tggctcctcgt	ctacatcaag	ggcctcttcc	cctcttcctc	ccagctctca	ggacaggtga	1260
ctgggagacc	ttgaaccctc	agcctcttcc	tttaaaaaaa	acaaaacaaa	acaaaactgt	1320
gggccattta	tttgggattt	tggagttgtt	tgggttttgt	ttttatatct	taatagttcg	1380
aaagtaagaa	gggagccctg	ctatggatgt	taagtccaaa	ttactcgggt	agtgggagca	1440
aaacctatga	cttccaaggg	gatgaggaga	ggttcagagg	acaggaggag	cctcccccat	1500
tgaaaaaaaaa	aaaatgggtc	aggacattcc	ctggatgagg	acaatgctag	gggtggcatc	1560
tcacatggct	gctgctattc	ctggtgcttc	cccacacttt	tgacagatgg	agtccttctc	1620
ctaccgcctc	ctgccacctc	accctacagg	cattctctat	gtaggaaaca	agagccttat	1680
cttatagagt	ggggagctga	gacacagctc	caggtaacac	tgacacagct	cccgaatgag	1740
gctgggacac	ctctgcaaac	tctcctcact	gtgctaaggg	tggcatgctc	ttgacaggaa	1800
acctaaatga	ccactcctct	catttggaaa	gtaatccact	gcagtaaaaag	tttcagacat	1860
gcaagagaga	gttttttttt	ttttttacta	caaatttttg	ctcccccata	aaattatttt	1920
tttattagag	ggagtatcca	agttttaaaa	gtatatagaa	ttttttgggt	gtaagagaaa	1980
tacatactca	ttaggatccc	gattaaattc	cttgagtaga	ctgggtgcct	ccagaaagca	2040
aagcaaagtt	aaacaaaacg	aaacaaaatc	cttcatatac	aaaaagaact	ttctgtttgt	2100
attggcagag	gtagttaggt	gattcaggta	ggctgaaaat	cctgggttgc	gggagcctca	2160
ctttattcca	ttcccacccg	ctttgatgtc	tatgcttggc	tctctgggct	gcccctggta	2220
ctgccgaatc	ctacacatct	cttatcagct	ttcctcaaac	tttaaggagg	ctctgtgagg	2280
gatgggtcat	gggaagaccc	aagctttccc	tccgccagga	ttgcaaaagc	aagtagactt	2340
ggtctatgca	gctcttcttc	caacaatttc	tttatttgga	attagaactt	cctttgtag	2400
tatctttgat	cttttgactc	aagcacattt	tgggaagggt	cccttacaaa	agtagaattt	2460
aaaacagagg	atacagttaa	agagcaaccc	aaaggacgct	taagaaaaccg	agaccacttc	2520
accgaacagg	actaaggaac	actttcgtgc	acagaagtca	gccgcaatcc	aggcacagga	2580
cgaagatggg	atacacgtgc	tcactctgtc	gtcctccttt	cctctccctc	cccagcgttc	2640
tagttagctt	gttgacttgt	taaaccttct	gttcttaaaa	tgaaaagcta	gcttacctca	2700
aagaatcttg	tttccattcg	gaaaccaacg	attttgtgtt	ttagaatgga	cagccctccc	2760
ctcaccactc	cctaccttgg	cctgggtgtc	ttgagacata	cggctcttgc	ttagtcgtgt	2820
gttggtctgt	ttgagcagga	acgaggcctc	caggccctga	ggtgggaagg	aaggattgga	2880
tgccactgcc	ctcctcccca	ctttagcatg	taggggccag	cccatctctt	ccagcagggt	2940
cctgctgagt	taccatagca	accagcaact	ccagggtacc	acaacagaca	atggctcagc	3000
gagccgacgt	gtggggatga	tgcaggggtt	ttggcccagc	cagaggaccc	agagttgagc	3060
ttcaaatgct	agagaagggg	agaaacagga	tgggaagggt	gtttaaggaa	ctggcagggg	3120
tctttgagtc	acatagagaa	gccgttgaag	gaggtagggc	aggttatctc	tgttccagtc	3180
accccccttc	agccccatcc	cacttctgtt	tcaaaactaa	gctcccacct	cgaacattga	3240
ccctttgtta	gaacaaagca	aagcatactc	ttaaacaaca	gtgttataaat	gagcctcaaa	3300
tgtatgtgga	tgagatctct	aagaagaggg	tcttctgggt	ttgattttta	aagaagagta	3360
tcctagtaaa	atatttaaaaa	aaaatttaaaa	agttttttaa	aaggaaacct	atgctattta	3420
aattggagcc	cagttgtaac	ttggttaaagg	caagcttctg	tacctttgtt	ataattaatt	3480
gtatacctgt	gtatgtaaat	ataaggcatt	cctattttgc	agttcagaac	aaaaaaaaact	3540
tatttgtaat	atagaataaa	gtttatttaa	aaataagaaa	aaaaaaaaaaa	aaaaaaa	3597

<210> 650

<211> 855

<212> DNA

<213> Homo sapiens

<400> 650

cgctcaatta	tctactcgag	tctagactcg	aggcggccgc	ccattgtgca	ctaaagcagg	60
ggatagcaac	ggcgctccctc	ctccccgctc	agctgcagcc	cgcagtcctc	acagtggtaa	120
catgccacgt	ggtagtctct	gtccatggac	accacacgga	tggttgctct	gcagccctgt	180
gcagggagga	taggacgggc	acaggaggcg	cattttggtg	caaaaaccgt	gtgatagtct	240
cgcacgcagt	agatgttggt	ctccacgtcc	acggtgaagg	gaaccccgct	caggcactca	300
ttgcacacgg	agcaccggaa	gcagcctggg	tggtaggact	tgcccagggc	ctgcaggatc	360
atttccatga	tgagatgtcc	acacacgctg	catttgctcg	ccgtctgctg	gaacccggag	420
tacaggaagt	cctcctggca	gtacaactttc	tcacccacgt	tgtagaacgc	cttcccacgg	480
agtcgtctcc	cacacgagtc	gcaggtgaag	cagtcagtgt	gataaagact	ccccattgcc	540
tggcacgcct	gctgggctcc	gtagatgcca	agcccacact	tgatgcaa	gccgaagtag	600
tcccgcgccg	tgcgcgctc	gagcgcccg	tccagctccc	gggtgagcgc	ctccagccgc	660
cgctcggccg	cgcttggggc	gccctcccg	ccagggggca	gcgggagtc	aggcagcggg	720
aagggggccg	gccccgcagg	ctccggggag	cgagcggg	cggggcaggc	gccgggcggg	780
aggaagtcag	cgtaggcttt	agtgcacaat	gggaattcgg	atccggtcga	cactagtctt	840
aggatctctt	ttttt					855

<210> 651

<211> 3791

<212> DNA

<213> Homo sapiens

<400> 651

acagacggcg	ggtgaacatg	gcgtcctcga	cttgggtctga	gacgtgatag	gcctgccttc	60
tggttgaaga	tgtggcgagt	gaaaaaactg	agcctcagcc	tgctgccttc	gccccagacg	120
ggaaaaccat	ctatgagaac	tcctctccgt	gaacttaccc	tgagcccg	tgccctcacc	180
acctctggaa	aaagatcccc	cgcttgetcc	tcgctgaccc	catcactgtg	caagctgggg	240
ctgcaggaag	gcagcaacaa	ctcgtctcca	gtggattttg	ttaaatacaa	gaggacagac	300
ttatcttcag	aacatttcag	tcattcctca	aagtggctag	aaacttgtca	gcataaatca	360
gatgagcagc	ctctagatcc	aattccccaa	attagctcta	ctcctaaaac	gtctgaggaa	420
gcagtagacc	cactgggcaa	ttatatgggt	aaaaccatcg	tccttgtagc	atctccactg	480
gggcagcaac	aagacatgat	atttgaggcc	cgtttagata	ccatggcaga	gacaaacagc	540
atatctttaa	atggaccttt	gagaacagac	gatctggtag	gagaggaggt	ggcaccttgc	600
atgggagaca	ggttttcaga	agttgctgct	gtatctgaga	aacctatctt	tcaggaatct	660
ccgtcccatc	tccttagagga	gtctccacca	aatccctggt	ctgaacaact	acattgtctc	720
aaggaaagcc	tgagcagtag	aactgaggct	gtgcgtgagg	acttagtagc	ttctgaaagt	780
aacgccttct	tgcttctctc	tgttctctgg	ctttccctt	caactgcctt	ggcagcagat	840
ttcctgtgtc	atcatgtgga	cccagaggag	gaaattgtag	agcatggagc	tatggaggaa	900
agagaaatga	ggtttccac	acatcctaag	gagcttgaaa	cagaagatca	agcacttgct	960
tcaagtgtgg	aagatattct	gtccacatgc	ctgacaccaa	atctagtaga	aatggaatcc	1020
caagaagctc	caggcccagc	agtagaagat	gttggttagga	ttcttggtct	tgatacagag	1080
tcttgatgt	ccccactggc	ctggctggaa	aaagggtgaa	atacctccgt	catgctggaa	1140
aatctccgcc	aaagcttatc	ccttccctcg	atgcttcggg	atgctgcaat	tggcactacc	1200
cctttctcta	cttgctcggt	ggggacttgg	tttactcctt	cagcaccaca	ggaaaagagt	1260
acaaacacat	cccagacagg	cctgggttgg	accaagcaca	gtacttctga	gacagagcag	1320
ctcctgtgtg	gccggcctcc	agatctgact	gccttgtctc	gacatgactt	ggaagataac	1380
ctgctgagct	ctcttgctcat	tgtggagttt	ctctcccgcc	agcttcggga	ctggaagagc	1440
cagctggctg	tccttcaccc	agaaaccag	gacagtagca	cacagactga	cacatctcac	1500
agtgggataa	ctaataaact	tcagcatctt	aaggagagcc	atgagatggg	acaggcccta	1560
cagcaggcca	gaaatgtcat	gcaatcatgg	gtgcttatct	ctaaagagct	gatatccttg	1620
cttcacctat	ccctgttgca	tttagaagaa	gataagacta	ctgtgaatca	ggagtctcgg	1680

cgtgcagaaa	cattggtctg	ttgctgtttt	gatttgctga	agaaattgag	ggcaaagctc	1740
cagagcctca	aagcagaaa	ggaggaggca	aggcacagag	aggaaatggc	tctcagaggg	1800
aaggatgcgg	cagagatagt	gttggaggct	ttctgtgcac	acgccagcca	gcgcatacag	1860
cagctggaac	aggacctagc	atccatgcgg	gaattcagag	gcctttctgaa	ggatgcccag	1920
acccaactgg	tagggcttca	tgccaagcaa	gaagagctgg	ttcagcagac	agtgagtcct	1980
acttctacct	tgcaacaaga	ctggagggtcc	atgcaactgg	attatacaac	atggacagct	2040
ttgctgagtc	gggtcccgaca	actcacagag	aaactcacag	tcaagagcca	gcaagccctg	2100
caggaacgtg	atgtggcaat	tgaggaaaag	caggagggttt	ctagggtgct	ggaacaagtc	2160
tctgcccagt	tagaggagtg	caaaggccaa	acagaacaac	tggagttgga	aaacattcgt	2220
ctagcaacag	atctccgggc	tcagttgcag	attctggcca	acatggacag	ccagctaaaa	2280
gagctacaga	gtcagcatat	ccattgtgcc	caggacctgg	ctatgaagga	tgagttactc	2340
tgccagctta	cccagagcaa	tgaggagcag	gctgctcaat	gcgtaaagga	agagatggca	2400
ctaaaacaca	tgcaggcaga	actgcagcag	caacaagctg	tcctggccaa	agaggtgcgg	2460
gacctgaaag	agaccttgga	gtttgcagac	caggagaatc	aggttgctca	cctggagctg	2520
ggtcagggtt	agtgtcaatt	gaaaaccaca	ctggaagtgc	tccgggagcg	cagcttgcat	2580
tgtgagaacc	tcaaggacac	tgtagagaac	ctaacggcta	aactggccag	caccatagca	2640
gataaccagg	agcaagatct	ggagaaaaca	cggcagtact	ctcaaaagct	agggctgctg	2700
actgagcaac	tacagagcct	gactctcttt	ctacagacaa	aactaaagga	gaagactgaa	2760
caagagaccc	ttctgctgag	tacagcctgt	cctcccaccc	aggaacaccc	tctgcctaata	2820
gacaggacct	tcctgggaag	catcttgaca	gcagtggcag	atgaagagcc	agaatcaact	2880
cctgtgccct	tgcttgggaag	tgacaagagt	gctttcaccc	gagtagcatc	aatggtttcc	2940
cttcagcccc	cagagacccc	aggcatggag	gagagcctgg	cagaaatgag	tattatgact	3000
actgagcttc	agagtcctttg	ttccctgcta	caagagtcta	aagaagaagc	catcaggact	3060
ctgcagcgaa	aaattttgtga	gctgcaagct	aggctgcagg	cccaggaaga	acagcatcag	3120
gaagtccaga	aggcaaaaga	agcagacata	gagaagctga	accaggcctt	gtgcttgccg	3180
tacaagaatg	aaaaggagct	ccagggaagt	atacagcaga	atgagaagat	cctagaacag	3240
atagacaaga	gtggcgagct	cataagcctt	agagaggagg	tgaccacact	taccgcctca	3300
cttcggcgctg	cggagacaga	gaccaaagt	ctccaggagg	cctggcaggc	cagctggact	3360
ccaactgcca	gcctatggcc	accaattgga	tccaggagaa	agtgtggctc	tctcaggagg	3420
tggacaaact	gagagtgatg	ttcctggaga	tgaaaaatga	gaaggaaaac	tcctgatcaa	3480
gttccagagc	ccatagaaat	atcctagagg	agaaccttcg	gcgctctgac	aaggagttag	3540
aaaaactaga	tgacattggt	cagcatattt	ataagaccct	gctctctatt	ccagagggtg	3600
tgaggggatg	caaagaacta	cagggattgc	tggaaatttct	gagctaagaa	actgaaagcc	3660
agaattttgt	tcacctcttt	ttacctgcaa	tacccctcta	ccccaatacc	aagaccaact	3720
ggcatagagc	caactgagat	aaatgctatt	taaataaagt	gtattttaatg	aaaaaaaaaa	3780
aaaaaaaaaa	a					3791

<210> 652

<211> 1683

<212> DNA

<213> Homo sapiens

<400> 652

ctctgagtg	ccagtgggtca	gttgccccag	gatgggggacc	acagccagag	cagccttggt	60
cttgacctat	ttggctgttg	cttctgctgc	ctctgaggga	ggcttcacgg	ctacaggaca	120
gaggcagctg	aggccagagc	actttcaaga	agttggctac	gcagctcccc	cctccccacc	180
cctatcccga	agcctcccca	tggatcaccc	tgactcctct	cagcatggcc	ctccctttga	240
gggacagagt	caagtgcagc	ccctccctc	tcaggaggcc	acccctctcc	aacaggaaaa	300
gctgctacct	gccaactcc	ctgctgaaaa	ggaagtgggt	ccccctctcc	ctcagggaagc	360
tgtccccctc	caaaaagagc	tgccctctct	ccagcaccct	aatgaacaga	aggaaggaac	420
gccagctcca	tttgggggacc	agagccatcc	agaacctgag	tcctggaatg	cagcccagca	480
ctgccaacag	gaccgggtccc	aaggggggctg	ggggccaccgg	ctggatggct	tcccccttgg	540
gcggccttct	ccagacaatc	tgaaccaaata	ctgccttctc	aaccgtcagc	atgtgggtata	600
tgggtccctg	aacctaccac	agtcacagta	ctcccacctc	actcgccagg	gtgagacctt	660

caatttcctg	gagattggat	attcccgcgtg	ctgccactgc	cgcagccaca	caaaccgcct	720
agagtgtgcc	aaacttgtgt	gggaggaagc	aatgagccga	ttctgtgagg	ccgagttctc	780
gggtcaagacc	cgacccact	gggtgtgcac	gcggcagggg	gaggctcggt	tctcctgctt	840
ccaggaggaa	gctccccagc	cacactacca	gctccgggcc	tgccccagcc	atcagcctga	900
tatttcctcg	ggctcttgagc	tgccctttccc	tcttgggggtg	cccacattgg	acaatatcaa	960
gaacatctgc	cacctgaggc	gcttccgcctc	tgtgccacgc	aacctgccag	ctactgacct	1020
cctacaaagg	gagctgtctg	cactgatcca	gctggagagg	gagttccagc	gctgctgccg	1080
ccaggggaac	aatcacacct	gtacatggaa	ggcctgggag	gatacccttg	acaataactg	1140
tgaccgggag	tatgtctgtga	agacccacca	ccacttgtgt	tgccgccacc	ctcccagccc	1200
tactcgggat	gagtgtcttg	cccgtcgggc	tccttaccctc	aactatgacc	gggacatctt	1260
gaccattgac	atcagtcgag	tcacccccaa	cctcatgggc	cacctctgtg	gaaaccaaag	1320
agttctcacc	aagcataaac	atattcctgg	gctgatccac	aacatgactg	cccgtgtctg	1380
tgacctgcc	tttccagaac	aggcctgtctg	tgagaggag	gagaaattaa	ccttcatcaa	1440
tgatctgtgt	gggtccccgac	gtaacatctg	gcgagaccct	gccctctgct	gttacctgag	1500
tcttggggat	gaacaggtca	actgcttcaa	catcaattat	ctgaggaacg	tggtcttagt	1560
gtctggagac	actgagaacg	ccaagggcca	gggggagcag	ggctcaactg	gaggaacaaa	1620
tatcagctcc	acctctgagc	ccaaggaaga	atgagtcacc	ccagagccct	agaggggtcag	1680
atg						1683

<210> 653

<211> 1696

<212> DNA

<213> Homo sapiens

<400> 653

cacctaaaag	ccaaaatggg	aaaggaaaag	actcatatca	acattgtcgt	cattggacac	60
gtagattcgg	gcaagtccac	cactactggc	catctgatct	ataaatgcgg	tggtcatcgac	120
aaaagaacca	ttgaaaaatt	tgagaaggag	gctgtgaga	tgaggaaagg	ctccttcaag	180
tatgcctggg	tcttgataa	actgaaagct	gagcgtgaac	gtggtatcac	cattgatatc	240
tccttgtgga	aatttgagac	cagcaagtac	tatgtgacta	tcattgatgc	cccaggacac	300
agagacttta	tcaaaaacat	gattacaggg	acatctcagg	ctgactgtgc	tgctcctgatt	360
gttgctgctg	gtgttggtga	atgtgaagct	ggatctctca	agaatgggca	gacccgagag	420
catgcccttc	tggtttacac	actgggtgtg	aaacaactaa	ttgtcgggtg	taacaaaatg	480
gattccactg	agccacccta	cagccagaag	agatatgagg	aaattgttaa	ggaagtccgc	540
acttacatta	agaaaatttg	ctacaacccc	gacacagtag	catttgtgcc	aatttctggt	600
tggaatggtg	acaacatgct	ggagccaagt	gctaacatgc	cttggttcaa	gggatggaaa	660
gtcaccgcga	aggatggcaa	tgccagtggg	accacgctgc	ttgaggctgt	ggactgcate	720
ctaccaccaa	ctcgtccaac	tgacaagccc	ttgcgcctgc	ctctccagga	tgtctacaaa	780
attggtggta	ttggtactgt	tcctgttggc	cgagtggaga	ctggtgttct	caaaccgggt	840
atggtgggtca	cctttgctcc	agtcaacggt	acaacgggag	taaaatctgt	cgaaatgcac	900
catgaagctt	tgagtgaagc	tcttcctggg	gacaatgtgg	gcttcaatgt	caagaatgtg	960
tctgtcaagg	atgttcgtcg	tggcaacggt	gctggtgaca	gcaaaaatga	cccaccaatg	1020
gaagcagctg	gcttcactgc	tcaggtgatt	atcctgaacc	atccaggcca	aataagcgcc	1080
ggctatgccc	ctgtattgga	ttgccacacg	gctcacattg	catgcaagtt	tgctgagctg	1140
aaggaaaaga	ttgatcgccg	ttctggtaaa	aagctggaag	atggccctaa	attcttgaag	1200
tctggtgatg	ctgccattgt	tgatatgggt	cctggcaagc	ccatgtgtgt	tgagagcttc	1260
tcagactatc	cacctttggg	tcgctttgct	gttcgtgata	tgagacagac	agttgcgggtg	1320
gggtgtcatca	aagcagtggg	caagaaggct	gctggagctg	gcaaggtcac	caagtctgcc	1380
cagaaagctc	agaaggctaa	atgaatatta	tccttaatac	ctgccacccc	actcttaatc	1440
agtgggtggaa	gaacgggtctc	agaactgttt	gtttcaattg	gccattttaag	tttagtagta	1500
aaagactgggt	taatgataac	aatgcctcgt	aaaaccttca	gaaggaaagg	agaatgtttt	1560
gtggaccact	ttggttttct	tttttgcgtg	tggcagtttt	aagttattag	tttttaaaat	1620
cagtactttt	taatggaaac	aacttgacca	aaaatttgtc	acagaatttt	gagaccatt	1680
aaaaaagtta	aatgag					1696

<210> 654
 <211> 7680
 <212> DNA
 <213> Homo sapiens

<400> 654
 gaagagcaag aggcaggctc agcaaattggt tcagccccag tccccggtgg ctgtcagtca 60
 aagcaagccc gggtgttatg acaatggaaa acactatcag ataaatcaac agtgggagcg 120
 gacctaccta ggtaattgtg tggtttgtac ttgttatgga ggaagccgag gttttaactg 180
 cgaaagtaaa cctgaagctg aagagacttg ctttgacaag tacactggga acacttaccg 240
 agtgggtgac acttatgagc gtccctaaaga ctccatgac tgggactgta cctgcatcgg 300
 ggctgggcca gggagaataa gctgtaccat cgcaaaccgc tgccatgaag ggggtcagtc 360
 ctacaagatt ggtgacacct ggaggagacc acatgagact ggtgggtaca tggtagagt 420
 tgtgtgtctt ggtaattggaa aaggagaatg gacctgcaag cccatagctg agaagtgttt 480
 tgatcatgct gctgggactt cctatgtggg cggagaaaac tgggagaagc cctaccaagg 540
 ctggatgatg gtagattgta cttgcctggg agaaggcagc ggacgcacac cttgcacttc 600
 tagaaataga tgcaacgatc aggacacaag gacatcctat agaattggag acacctggag 660
 caagaaggat aatcgaggaa acctgctcca gtgcatctgc acaggcaacg gccgaggaga 720
 gtggaagtgt gagaggcaca cctctgtgca gaccacatcg agcggatctg gccccttcac 780
 cgatgttcgt gcagctgttt accaaccgca gcctcaccac cagcctcctc cctatggcca 840
 ctgtgtcaca gacagtgggt tgggtctactc tgtggggatg cagtgggtga agacacaagg 900
 aaataagcaa atgctttgca cgtgcctggg caacggagtc agctgccaaag agacagctgt 960
 aaccagact tacggtggca acttaaattg agagccatgt gtcttaccat tcacctacaa 1020
 tggcaggacg ttctactcct gcaccacgga agggcgacag gacggacatc tttggtgcag 1080
 cacaacttcg aattatgagc aggaccagaa atactctttc tgacagacc acactgtttt 1140
 ggttcagact caaggaggaa attccaatgg tgcttctgac cacttccctc tcctatacaa 1200
 caaccacaat tacactgatt gcacttctga gggcagaaga gacaacatga agtgggtgtg 1260
 gaccacacag aactatgatg ccgaccagaa gtttggttcc tgcccatggc ctgcccacga 1320
 ggaaatctgc acaaccaatg aaggggtcat gtaccgcatt ggagatcagt gggataagca 1380
 gcatgacatg ggtcacatga tgaggtgcac gtgtgttggg aatggctcgt gggaaatggac 1440
 atgcattgcc tactcgcaac ttcgagatca gtgcattgtt gatgacatca cttacaatgt 1500
 gaacgacaca ttccacaagc gtcacgaaga ggggcacatg ctgaactgta catgcttcgg 1560
 tcaggggtcgg ggcagggtgga agtgtgatcc cgtcgaccaa tgccaggatt cagagactgg 1620
 gacgttttat caaattggag attcatggga gaagtatgtg catggtgtca gataccagtg 1680
 ctactgctat ggccgtggca ttggggagtg gcattgccaa cctttacaga cctatccaag 1740
 ctcaagtggc cctgtcgaag tatttatcac tgagactccg agtcagccca actcccaccc 1800
 catccagtgg aatgcaccac agccatctca catttccaag tacattctca ggtggagacc 1860
 taataattct gtaggcggtt ggaaggaagc taccatacca ggccacttaa actcctacac 1920
 catcaaaggc ctgaagcctg gtgtggtata cgaggggcag ctcacagca tccagcagta 1980
 cggccaccaa gaagtgactc gctttgactt caccaccacc agcaccagca cactgtgac 2040
 cagcaacacc gtgacaggag agacgactcc cttttctcct cttgtggcca cttctgaatc 2100
 tgtgaccgaa atcacagcca gtagctttgt ggtctcctgg gtctcagctt ccgacaccgt 2160
 gtcgggattc cgggtggaat atgagctgag tgaggaggga gatgagccac agtacctgga 2220
 tcttccaagc acagccactt ctgtgaacat ccctgacctg cttcctggcc gaaaatacat 2280
 tgtaaatgtc tatcagatat ctgaggatgg ggagcagagt ttgactctgt ctacttcaca 2340
 aacaacagcg cctgatgccc ctctgacccc gactgtggac caagttgatg acacctcaat 2400
 tgttgttcgc tggagcagac cccaggctcc catcacaggg tacagaatag tctattcgcc 2460
 atcagtagaa ggtagcagca cagaactcaa ccttctgaa actgcaact ccgtcaccct 2520
 cagtgacttg caacctggtg ttcagtataa catcactatc tatgctgtgg aagaaaatca 2580
 agaaagtaca cctgttgtca ttcaacaaga aaccactggc accccacgct cagatacagt 2640
 gccctctccc agggacctgc agtttgtgga agtgacagac gtgaagggtc ccatcatgtg 2700
 gacaccgcct gagagtgcag tgaccggcta ccgtgtggat gtgatccccg tcaacctgcc 2760
 tggcgagcac gggcagaggc tgcccatcag caggaacacc tttgcagaag tcaaccgggt 2820
 gtcccctggg gtcacctatt acttcaaagt ctttgagtg agccatggga gggagagcaa 2880

gcctctgact	gctcaacaga	caaccaaact	ggatgctccc	actaacctcc	agtttgtcaa	2940
tgaaactgat	tctactgtcc	tggtgagatg	gactccacct	cgggcccaga	taacaggata	3000
ccgactgacc	gtgggcctta	cccgaagagg	ccagcccagg	cagtacaatg	tgggtccctc	3060
tgtctccaag	tacccctga	ggaatctgca	gcctgcatct	gagtacaccg	tatccctcgt	3120
ggccataaag	ggcaaccaag	agagcccca	agccactgga	gtctttacca	cactgcagcc	3180
tgggagctct	attccacctt	acaacaccga	ggtgactgag	accaccatcg	tgatcacatg	3240
gacgcctgct	ccaagaattg	gttttaagct	gggtgtacga	ccaagccagg	gaggagaggc	3300
accacgagaa	gtgacttcag	actcaggaag	catcgttgtg	tccggcttga	ctccaggagt	3360
agaatacgtc	tacaccatcc	aagtcctgag	agatggacag	gaaagagatg	cgccaattgt	3420
aaacaaagtg	gtgacaccat	tgtctccacc	aacaaacttg	catctggagg	caaaccctga	3480
cactggagtg	ctcacagtct	cctgggagag	gagcaccacc	ccagacatta	ctgggttatag	3540
aattaccaca	accctacaa	acggccagca	gggaaattct	ttggaagaag	tggtccatgc	3600
tgatcagagc	tcttgcaactt	ttgataacct	gagtcceggc	ctggagtaca	atgtcagtgt	3660
ttacactgtc	aaggatgaca	aggaaagtgt	ccctatctct	gataccatca	tcccagctgt	3720
tctctctccc	actgacctgc	gattcaccaa	cattggtcca	gacaccatgc	gtgtcacctg	3780
ggctccaccc	ccatccattg	atttaaccaa	cttccctggg	cgttactcac	ctgtgaaaaa	3840
tgaggaagat	gttgacagagt	tgtcaatttc	tccttcagac	aatgcagtgg	tcttaacaaa	3900
tctctctgct	ggtacagaat	atgtagtgag	tgtctccagt	gtctacgaac	aacatgagag	3960
cacacctctt	agaggaagac	agaaaacagg	tcttgattcc	ccaactggca	ttgacttttc	4020
tgatattact	gccaactctt	ttactgtgca	ctggattgct	cctcgagcca	ccatcactgg	4080
ctacaggatc	cgccatcatc	ccgagcactt	cagtgaggaga	cctcgagaag	atcgggtgcc	4140
ccactctcgg	aattccatca	ccctcaccaa	cctcactcca	ggcacagagt	atgtggtcag	4200
catcgttgct	cttaatggca	gagaggaaag	tcctttattg	attggccaac	aatcaacagt	4260
ttctgatgtt	ccgagggacc	tggaagtgtg	tgctgcgacc	cccaccagcc	tactgatcag	4320
ctgggatgct	cctgctgtca	cagttagata	ttacaggatc	acttacggag	aaacaggagg	4380
aaatagccct	gtccaggagt	tcactgtgcc	tgggagcaag	tctacagcta	ccatcagcgg	4440
ccttaaacct	ggagttgatt	ataccatcac	tgtgtatgct	gtcactggcc	gtggagacag	4500
ccccgcaagc	agcaagccaa	tttccattaa	ttaccgaaca	gaaattgaca	aaccatccca	4560
gatgcaagtg	accgatgttc	aggacaacag	cattagtgtc	aagtggctgc	cttcaagttc	4620
ccctgttact	ggttacagag	taaccaccac	tcccaaaaat	ggaccaggac	caacaaaaac	4680
taaaactgca	ggtccagatc	aaacagaaat	gactattgaa	ggcttgccagc	ccacagtgga	4740
gtatgtgggt	agtgtctatg	ctcagaatcc	aagcggagag	agtcagcctc	tggttcagac	4800
tgcagtaacc	aacattgato	gccctaaagg	actggcattc	actgatgtgg	atgtcgattc	4860
catcaaaatt	gcttgggaaa	gcccacaggg	gcaagtttcc	aggtacaggg	tgacctactc	4920
gagccctgag	gatggaatcc	atgagctatt	ccctgcacct	gatggtgaag	aagacactgc	4980
agagctgcaa	ggcctcagac	cgggttctga	gtacacagtc	agtgtggttg	ccttgccagc	5040
tgatatggag	agccagcccc	tgattggaac	ccagtcacac	gctattcctg	caccaactga	5100
cctgaagttc	actcaggtca	caccacaaag	cctgagcgcc	cagtggacac	cacccaatgt	5160
tcagctcact	ggatatcgag	tgcgggtgac	ccccaggag	aagaccggac	caatgaaaga	5220
aatcaacctt	gctcctgaca	gctcatccgt	ggttgtatca	ggacttatgg	tggccaccaa	5280
atatgaagtg	agtgtctatg	ctcttaagga	cactttgaca	agcagaccag	ctcaggggtg	5340
tgtcaccact	ctggagaatg	tcagocccacc	aagaagggct	cgtgtgacag	atgtctactga	5400
gaccaccatc	accattagct	ggagaaccaa	gactgagacg	atcactggct	tccaagttga	5460
tgccgttcca	gccaatggcc	agactccaat	ccagagaacc	atcaagccag	atgtcagaag	5520
ctacaccatc	acagggtttac	aaccaggcac	tgactacaag	atctacctgt	acaccttgaa	5580
tgacaatgct	cggagctccc	ctgtggtcat	cgacgcctcc	actgccattg	atgcaccatc	5640
caacctgctg	ttcctggcca	ccacacccaa	ttccttgctg	gtatcatggc	agccgccacg	5700
tgccaggatt	accggctaca	tcatacaagta	tgagaagcct	gggtctcctc	ccagagaagt	5760
ggtccctcgg	ccccgccctg	gtgtcacaga	ggctactatt	actggcctgg	aaccgggaac	5820
cgaatatata	atttatgtca	ttgccctgaa	gaataatcag	aagagcgagc	ccctgattgg	5880
aaggaaaaag	acagacgagc	ttccccaaact	ggtaaccctt	ccacacccca	atcttcatgg	5940
accagagatc	ttggatgttc	cttccacagt	tcaaaagacc	cctttcgtca	cccaccctgg	6000
gtatgacact	ggaaatggta	ttcagcttcc	tggcacttct	ggtcagcaac	ccagtgttgg	6060
gcaacaaatg	atctttgagg	aacatggttt	taggcggacc	acaccgccc	caacggccac	6120

ccccataagg	cataggccaa	gaccataccc	gccgaatgta	ggacaagaag	ctctctctca	6180
gacaaccatc	tcatgggccc	cattccagga	cacttctgag	tacatcattt	catgtcatcc	6240
tggtggcact	gatgaagaac	cettacagtt	cagggttcc	ggaacttcta	ccagtgccac	6300
tctgacaggc	ctcaccagag	gtgccaccta	caacatcata	gtggaggcac	tgaaagacca	6360
gcagaggcat	aagggttcggg	aagaggttgt	taccgtgggc	aactctgtca	acgaaggctt	6420
gaaccaacct	acggatgact	ctgtccttga	cccctacaca	gtttcccatt	atgccgttgg	6480
agatgagtgg	gaacgaatgt	ctgaatcagg	ctttaaactg	ttgtgccagt	gcttaggctt	6540
tggaagtgg	catttcagat	gtgattcatc	tagatggtgc	catgacaatg	gtgtgaacta	6600
caagattgga	gagaagtggg	accgtcaggg	agaaaatggc	cagatgatga	gctgcacatg	6660
tcttggaac	ggaaaaggag	aattcaagt	tgacctcat	gaggcaacgt	gttacgatga	6720
tggaagaca	taccacgtag	gagaacagt	gcagaaggaa	tatctcgggtg	ccatttgctc	6780
ctgcacatgc	tttgagggcc	agcggggctg	gcgctgtgac	aactgccgca	gacctggggg	6840
tgaaccagc	cccgaaggca	ctactggcca	gtcctacaac	cagtattctc	agagatacca	6900
tcagagaaca	aacactaatg	ttaattgccc	aattgagtgc	ttcatgcctt	tagatgtaca	6960
ggctgacaga	gaagattccc	gagagtaaat	catctttcca	atccagagga	acaagcatgt	7020
ctctctgcca	agatccatct	aaactggagt	gatgttagca	gacctagctt	agagttcttc	7080
tttctttctt	aagccctttg	ctctggagga	agttctccag	cttcagctca	actcacagct	7140
tctccaagca	tcaccctggg	agtttccctga	gggttttctc	ataaatgagg	gctgcacatt	7200
gcctgttctg	cttcgaagta	ttcaataccg	ctcagttatt	taaatgaagt	gattctaaga	7260
tttggttttg	gatcaatagg	aaagcatatg	cagccaacca	agatgcaaat	gttttgaaat	7320
gatatgacca	aaattttaag	taggaaagtc	accctaacac	ttctgctttc	acttaagtgt	7380
ctggcccgc	atactgtagg	aacaagcatg	atcttggttac	tgtgatattt	taaatatcca	7440
cagtactcac	tttttccaaa	tgatcctagt	aattgcctag	aaatatcttt	ctcttacctg	7500
ttatattatca	atttttccca	gtatttttat	acggaaaaaa	ttgtattgaa	aacacttagt	7560
atgcagttga	taagaggaat	ttggtataat	tatggtgggt	gattattttt	tatactgtat	7620
gtgccaaagc	tttactactg	tggaaagaca	actgttttaa	taaaagattt	acattccaca	7680

<210> 655

<211> 1316

<212> DNA

<213> Homo sapiens

<400> 655

tcctaatacg	actcactata	gggctcgagc	ggccgcccgg	gcaggctcga	tgcaggcgac	60
ttgcgagctg	ggagcgattt	aaaacgcttt	ggattccccc	ggcctgggtg	gggagagcga	120
gctgggtgcc	ccctagattc	cccgcctccg	cacctcatga	gccgacctc	ggctccatgg	180
agcccggcaa	ttatgccacc	ttggatggag	ccaaggatat	cgaaggcttg	ctgggagcgg	240
gagggggg	gaatctgggt	gcccactccc	ctctgaccag	ccacctcagc	gcgcctacgc	300
tgatgcctgc	tgtaactat	gcccccttgg	atctgccagg	ctcggcggag	ccgccaaagc	360
aatgccaccc	atgccttggg	gtgcccagg	ggacgtcccc	agctcccgtg	ccttatgggt	420
actttggagg	cgggtactac	tcctgcccag	tgtcccggag	ctcgtgaaa	ccctgtgccc	480
aggcagccac	cctggccgcg	taccgcgcgg	agactccac	ggccggggaa	gagtacccca	540
gtcgccccac	tgagtttgcc	ttctatccgg	gatateccgg	aacctaccac	gctatggcca	600
gttacctgga	cgtgtctgtg	gtgcagactc	tgggtgtccc	tggagaaccg	cgacatgact	660
ccctgttgcc	tgtggacagt	taccagtctt	gggctctcgc	tgggtggctg	aacagccaga	720
tgtgttgcca	gggagaacag	aaccaccag	gtcccctttg	gaaggcagca	tttgagact	780
ccagcgggca	gcacctcct	gacgcctgcg	cctttcgtcg	cggccgcaag	aaacgcattc	840
cgtacagcaa	ggggcagttg	cgggagctgg	agcgggagta	tgcggctaac	aagtccatca	900
ccaaggacaa	gaggcgcaag	atctcggcag	ccaccagcct	ctcggagcgc	cagattacca	960
tctgggttca	gaaccgcggg	gtcaaagaga	agaaggttct	cgccaagggtg	agaacagcgg	1020
ctaccoccta	agagatctcc	ttgcctgggt	gggaggagcg	aaagtggggg	tgctcctgggg	1080
agaccagaaa	cctgccaaagc	ccaggctggg	gccaaaggact	ctgctgagag	gcccctagag	1140
acaacacct	tcccaggcca	ctggctgctg	gactgttctt	caggagcggc	ctgggtaccc	1200

agtatgtgca	gggagacgga	accccatgtg	acaggccccc	tccaccaggg	ttcccaaaga	1260
acctggccca	gtcataatca	ttcatcctca	cagtggcaat	aatcacgata	accagt	1316

<210> 656

<211> 566

<212> DNA

<213> Homo sapiens

<400> 656

cccaccaaac	ccataaagag	ggtgggtcga	cccacgcgtc	cgcggacgcg	tgggaaatta	60
ttgaattgga	aacagaaata	gaaaagtitta	aagctgagaa	cgcattcttta	gctaaaacttc	120
gcattgaacg	agaaagtgcc	ttggaaaaac	tcaggaaaga	aattgcagac	ttcgaacaac	180
agaaagcaaa	agaattagct	cgaatagaag	agtttaaaaa	ggaggagatg	aggaagctac	240
aaaaggaacg	taaagttttt	gaaaagtata	ctacagctgc	aagaactttt	ccagataaaa	300
aggaacgtga	agaaatacac	actttaaaac	agcaaatagc	agatttacgg	gaagatttga	360
aaagaaagga	gaccaaattg	tcaagtacac	acagccgtct	cagaagccag	atacaaatgt	420
tagtcagaga	gaacacagac	ctccgggaag	aaataaaaagt	gatggaaaga	ttccgactgg	480
atgcctggaa	gagagcagaa	gccatagaga	gcagcctcga	ggtggagaag	aaggacaagc	540
ttgcgaacac	atctgttcga	tttcaa				566

<210> 657

<211> 5067

<212> DNA

<213> Homo sapiens

<400> 657

gccccgacac	ctgtctgcag	catggataag	tatgacgacc	tgggcctgga	ggccagtaaa	60
ttcatcgagg	acctgaacat	gtatgaggcc	tctaaggatg	ggctcttccg	agtggacaag	120
ggtgaggcca	acaaccccg	gtttgaggaa	actcgcaggg	tgttcgccac	caagatggcc	180
aaaatccacc	tccagcagca	gcagcagcag	ctcctgcagg	aggagactct	gcccaggggg	240
agtagaggcc	ctgtcaatgg	agggggccgc	ctggggccac	aggcccgttg	ggaagtgtg	300
ggcagcaagc	tgactgtgga	tgggtgtgcc	aagcctcctc	ttgctgcctc	gacaggggca	360
cctggggcag	tcaccaccct	cgtgctggg	cagccccgt	acccaccgca	ggagcagaga	420
tccaggccat	acctgcatgg	cacgaggcat	ggcagccagg	actgtggttc	cagggagagc	480
ctggcgactt	ctgagatgtc	tgctttccac	cagccaggcc	cctgtgagga	tccttccctgc	540
ctcactcatg	gagactatta	tgacaacctc	tccttggcaa	gccccaaagt	gggtgacaaa	600
ccaggagtgt	ccccagcat	cggcctgagt	gtagggagtg	ggtggcctag	ctccccggg	660
agtgacccac	cactgcccac	accctgcggg	gacctcccc	taaatcaccg	acagctctcc	720
ctgagctcca	gcaggctctc	tgagggtagc	ctcgggtggtc	agaatagtgg	cattgggtggc	780
cgcagcagcg	agaagccaac	aggcctttgg	tccactgcct	cctcccagcg	ggtgagccct	840
ggcctgcctt	ccccaaactt	ggagaacgga	gcaccagctg	tggggcctgt	tcagcccagg	900
actccttctg	tgtcagcacc	cttggccctg	agctgcccc	ggcaaggagg	tcttccaaga	960
tcaaactcgg	ggctgggggg	tgaggtttca	ggtgtgatgt	ccaaacccaa	tgtggacccc	1020
caaccctggg	tccaggatgg	gccccaatct	tacctttcca	gttctgcccc	gtcatcctcg	1080
ccagctggcc	tggacgggtc	acagcagggt	gcggctccctg	ggctggggcc	gaagcctggc	1140
tgcacagacc	ttggcactgg	tcccaagctc	agccccacca	gtcttgcca	tccagtgatg	1200
tccaccctgc	ctgagttatc	ttgtaaagag	ggtcccctgg	gctgggtctc	tgatggtagc	1260
ctgggatctg	tgctcctgga	cagccccagc	tcccctaggg	taaggctgcc	ctgccagccc	1320
ctcgtcccag	gtcctgagct	gagaccctct	gctgctgagt	tgaaattaga	agccctcacc	1380
caacgtctgg	agcgagagat	ggatgctcac	ccgaaggctg	attactttgg	agcctgtgtg	1440
aatgcagca	aaggggtgtt	tggggctggc	caggcctgtc	aggccatggg	gaacctctac	1500
catgacacat	gcttcacctg	tgcagcttgc	agccgggaagc	tgagaggaaa	agccttttat	1560
tttgtcaacg	gcaaagtgtt	ttgtgaagaa	gacttcctgt	actctggttt	ccagcagtcg	1620
gctgacaggt	gttttctttg	tggacatctg	atcatggaca	tgatcctgca	agccctgggg	1680

aagtcctacc	accccggtg	tttccgctgt	gtcatctgta	atgagtgttt	ggatggggtg	1740
cccttcaccg	tggactcaga	gaacaagatc	tactgtgtcc	gagattacca	caagggtgctg	1800
gcccccaagt	gtgcagcctg	tgggcttccc	atccttccac	ctgagggtc	agatgagacc	1860
atccgtgtcg	tgtccatgga	cagagactac	cacgtggagt	gttaccactg	cgaggactgt	1920
ggtctggagc	tcaatgatga	agatggccac	cgctgttatc	cgctggagga	ccacctgttc	1980
tgtcactcct	gccacgtgaa	gaggctggag	aagagaccct	catctacagc	ccttcaccag	2040
caccacttct	agccagagcc	acttgacagc	atcacggcag	gggatgagga	gccggggttg	2100
ctgctgtctg	ttccggtggc	ccctgggggtg	gaagtggggt	aggggaagag	gaggggcagg	2160
agggagagtt	cctgtgagca	tgtggggggt	gcctttcctt	taaccaggga	ggtgaacact	2220
acctgcctcc	tgcgtgtatt	ttccaagtgc	ttttctctgt	tgccacattt	tcctcagggt	2280
actcaggaaa	atgctccagc	atgtgcgagc	acatgacctg	aggttgcatc	atagcaccaa	2340
aggaatcctc	ctgtcccctc	tgggaacatt	tcatgcttca	gaggagagag	tttttattga	2400
gcttgtttca	caatatcccc	ttgaagggtg	agctcagctg	ccaatacatt	caaccctttc	2460
tcttccttca	gaaaaatacc	tatacccaaa	tgttccctcc	cccgcataat	atcatggcat	2520
gacttaaggc	ttctttttcac	ctgagagctt	cagttcttct	gcagaatggc	tgcaaattta	2580
attgcattaa	ggcaagaagg	aagctctaata	gtgtgctttg	tatcctaaga	taaattttgt	2640
tagaaaacca	gagtcaagat	ttgaaatagg	tgaggcaggg	tttctcctt	agacactgac	2700
agcattctcc	gtaccccttc	aaatccttac	tctcctaaag	gcagctgagt	ccgcgacaga	2760
aatttgccct	atgggagtaa	aacatacttt	gggagaagaa	cttggtgcag	gcaccaggat	2820
tttttttttt	gccacgtgt	ttgcgctgtt	tttctctgga	gttctcaaga	gttggtgact	2880
tggaaaggcg	cttctgcaag	gcaagtctca	ggaacccatg	caggtaacatc	gcttgacact	2940
gttttttagct	tattttaatga	cgggcttttg	ggaagagctg	cccgcatact	gagagacagc	3000
ttcttataaa	caaggagagt	ttttgtgtgt	gcgagatctc	taagccagcg	tgggagggag	3060
cgctcagga	taagtattta	tattcatttc	gttggtttct	ctcctgcccc	attcctggca	3120
caggcattat	gtttgaagaa	accaggataa	ggtacactgc	ttttgtctgt	ttaatTTTTT	3180
tagttgtttc	ccttcacttt	cagtcttcca	cacacaaaaa	atacctcaca	gagcttcacc	3240
aatcacaga	ttcaggagga	atttggtctt	cacactggac	tcagatacct	tcttcagtgt	3300
gttggaatc	actggcttca	cacaggccca	actccaactg	gtcagggcag	agtgatcgta	3360
actaaaggtc	agtggggaat	agatccgatt	cagtgtcttt	gccttatgca	tttcagcatc	3420
ctggctcccc	agggtggcag	gagctgagga	agggccacac	actggcaaga	tttcaagacc	3480
actctctgca	ctgaagaggt	aaaatttgca	ctgcaagtca	catccctgag	gccagagggtc	3540
agtacccttt	ggtatttcga	ttagaagaag	ctgcaaaaaga	aaggcagccc	attttaccat	3600
tgccagccag	gccggggaca	caggagccgg	tgtgtgcact	ctgcctccta	acattgcacc	3660
cagagcaaga	ggactgggtg	ctgggctgca	gaggccgggtc	agtggagccc	ctagcacgtg	3720
tgaactcagg	cttttcattg	ggcccggctc	cacttctagg	ccatgttttg	actcatttgg	3780
taaccattgc	ctgtaagcag	cacagaattg	gtgccatgga	ttatcttttc	catgttgatg	3840
gaattcattc	tgttggaatc	ctttggccag	atgtcacttc	agccagggtg	tgcacatca	3900
ttggttcttt	ttcacaggct	gagcctcctg	aaaacccatg	aacgctgggg	ctgggggaagt	3960
gaaccctgag	gtggggacc	tctcttccca	tcaaatcatc	cagctcagtg	tggggcggtg	4020
caggggggta	aatgaagcca	gccaatgtgt	taacctgtct	ctgtcaacct	aagaatgttg	4080
gccttactga	cacacctttg	ctccatgttc	aagaccagaa	gtagctggga	tttgtttgca	4140
aattgggtaa	ttagtttaaa	aatctgtgat	tacattttta	aatgaaattt	tcaaagtggc	4200
ctagattgag	gtgattcaga	taggtttgctg	aatataccat	tttatattgt	tgagaaagaa	4260
caaaaaggga	atttccagat	gtcctagaaa	tcctagcaac	agattttctct	ggttgtcagt	4320
ttccctggag	aaggcgccag	ataggaatct	ccaatcagtt	gtttttctct	tcgcttcagg	4380
cccttacaca	aaagccatga	agagatgttc	acctaccggg	tatttttaa	gttctgtaaa	4440
ttattagcca	aatagaactg	taatgggggt	gtatttatgg	gcgcctagaa	agaaaacaca	4500
aggacttgg	aggccaggaa	gaaaagattt	taaaatttag	aatgaatagc	ccttctgggt	4560
tttctttttg	acaattcttg	gacttgaggt	aaaacaagga	ggattgtggc	cggatttcag	4620
atcccaaagc	cagcctccat	cttaggcctt	tgccctcattg	tgcccttttag	gttttcttac	4680
ccaccgtctc	ctgtttttgtc	ttttttttct	tttctcctac	ccctatcttg	ggacattcag	4740
aaactgcctg	ggtggtttga	gaagagacaa	cccagtttga	tctgcaatac	aaggatccat	4800
tcgtaatctc	tctctcactg	atgttattcc	cccatctgcc	gtcttggttc	atctcaccac	4860
agaagggtcat	ttagtctctac	ccagccatcg	gctgcgggtat	gacagcagga	tggcacttcc	4920

cattttctctg	tggttagtgc	tcgagtga	aaacctcttca	gctgagtcct	ctgagggttct	4980
gctgttgagt	cctgggtggc	tgatggaatg	attgaggagg	tctgggtcacc	ctcaagcgcc	5040
gtcatcgct	tgtttccatg	ggcttct				5067

<210> 658

<211> 950

<212> DNA

<213> Homo sapiens

<400> 658

tcgaccggat	ccgaattccc	attgtgcact	aaagcgtctc	cctgctccgc	ggccccgggct	60
ggcggggcggg	cgctcgggtg	gcggctgcag	cagcagaggg	agacccgcgg	caaccccgcc	120
aacccagggc	tcggcgctgc	tgccaccatg	acgggaagca	atatgtcggg	cgccctggcc	180
aacgccgtgt	gccagcgctg	ccaggccccg	ttctcccccg	ccgagcgcat	tgtcaacagc	240
aatggggagc	tgtaccatga	gcactgcttc	gtgtgtgccc	agtgtctccg	gcccttcccc	300
gaggggctct	tctatgagtt	tgaaggccgg	aagtactgcg	aacacgactt	ccaaatgctg	360
tttgctccgt	gctgtggatc	ctgcgggtgag	ttcatcattg	gccgcgtcat	caaggccatg	420
aacaacaact	ggcacccggg	ctgcttccgc	tgcgagctgt	gtgatgtgga	gctggctgac	480
ctgggctttg	tgaagaatgc	cggcaggcat	ctctgccggc	cttgccacaa	ccgtgagaag	540
gccaagggcc	tgggcaagta	catctgccag	cgggtgccacc	tggtcacga	cgagcagccc	600
ctcatgtttc	ggagcgacgc	ctaccacctc	gaccacttca	actgcaccca	ctgtgggaag	660
gagctgacag	ccgaggcccg	cgagctgaag	ggtgagctct	actgcctgcc	ctgccatgac	720
aagatgggcg	ttcccatctg	cggggcctgc	cgccggccca	tcgagggccg	agtgggtcaac	780
gcgctgggca	agcagtggca	cgtggagcac	tttgtctgtg	ccaagtgtga	gaagccattc	840
ctggggcacc	ggcactatga	gaagaagggc	ctggcctact	gcgagcttta	gtgcacaatg	900
ggcggccgcc	tcgagtctag	actcgagtag	ataattgagc	ggaatttctt		950

<210> 659

<211> 2346

<212> DNA

<213> Homo sapiens

<400> 659

ccgcgctcgc	ccccgcctcc	ccctgectca	gcggctgccc	ccgccagcgg	gccgcccgcct	60
ccccggggcc	ttgcagcggg	ccccggcccc	gctggagggg	ccccgacccc	agctctgggtg	120
gcgggcagca	gcgcgcgggc	ccccctccct	cacggggact	cgccctgaa	cgagcaggag	180
aaggagttgc	agcggcggtc	gaagcgtctc	taccgggccg	tggacgaaca	agagacgccg	240
ctgcctcggt	cctggagccc	gaaggacaag	ttcagctaca	tcggcctctc	tcagaacaac	300
ctgcgggtgc	actacaaaag	tcattggcaa	accccaaaag	atgccgcgtc	agttcgagcc	360
acgcatccaa	taccagcagc	ctgtgggatt	tattattttg	aagtataaat	tgtcagtaag	420
ggaagagatg	gttacatggg	aattgggtctt	tctgctcaag	gtgtgaacat	gaatagacta	480
ccagggttggg	ataagcattc	atatgggttac	catggggatg	atggacattc	gttttgttct	540
tctggaactg	gacaacctta	tggaaccaact	ttcactactg	gtgatgtcat	tggtgtgtgt	600
gttaatctta	tcaacaatac	ctgctttttac	accaagaatg	gacatagttt	aggtattgct	660
ttcactgacc	taccgccaac	tttgtatcct	actgtggggc	ttcaaacacc	aggagaagtg	720
gtcgatgcc	atgttgggca	acatcctttc	gtgtttgata	tagaagacta	tatgcgggag	780
tggagaacca	aaatccaggc	acagatagat	cgatttccta	tcggagatcg	agaaggagaa	840
tggcagacca	tgatacaaaa	aatgggtttc	tcttatttag	tccaccatgg	gtactgtgcc	900
acagcagagg	cctttgccag	atctacagac	cagaccgttc	tagaagaatt	agcttccatt	960
aagaatagac	aaagaattca	gaaattggta	ttagcaggaa	gaatgggaga	agccattgaa	1020
acaacacaac	agttataccc	aagtttactt	gaaagaaatc	ctaattctcct	tttcacatta	1080
aaagtgcgtc	agtttataga	aatgggtgaat	ggtacagata	gtgaagtacg	atgtttggga	1140
ggccgaagtc	caaagtctca	agacagttat	cctgttagtc	ctcgaccttt	tagtagtcca	1200

agtatgagcc	ccagccatgg	aatgaatata	cacaatttag	catcaggcaa	aggaagcacc	1260
gcacattttt	cagggttttga	aagttgtagt	aatggtgtaa	tatcaaataa	agcacatcaa	1320
tcatattgcc	atagtaataa	acaccagtca	tccaacttga	atgtaccaga	actaaacagt	1380
ataaatatgt	caagatcaca	gcaagttaat	aacttcacca	gtaatgatgt	agacatggaa	1440
acagatcact	actccaatgg	agttggagaa	acttcatcca	atggtttctt	aaatggtagc	1500
tctaaacatg	accacgaaat	ggaagattgt	gacaccgaaa	tggaagttga	ttcaagtcag	1560
ttgagacgcc	agttgtgtgg	aggaagtcag	gccgccatag	aaagaatgat	ccactttgga	1620
cgagagctgc	aagcaatgag	tgaacagcta	aggagagact	gtggcaagaa	cactgcaaac	1680
aaaaaaatgt	tgaaggatgc	attcagtcta	ctagcatatt	cagatccctg	gaacagccca	1740
gttggaatc	agcttgaccc	gattcagaga	gaacctgtgt	gctcagctct	taacagtgca	1800
atattagaaa	cccacaatct	gccaaagcaa	cctccacttg	ccctagcaat	gggacaggcc	1860
acacaatgtc	taggactgat	ggctcgatca	ggaattggat	cctgcgcatt	tgccacagtg	1920
gaagactacc	tacattagct	atgcatttca	agagctcaca	cttatattgt	ggcatatagt	1980
caacatggaa	gtagaccagc	tctgctgatt	tgaaatttag	atTTTTTtaa	ttatgtactg	2040
gggacagggt	tttgtcgctt	tacattgctt	cctagtttac	agcatgatgc	aaatgatttt	2100
ctaacttagt	gttaggagaa	attattttcc	atctttaaacc	tcttagttgt	ctaagagtta	2160
aatattactg	aatttcagac	gttcaaattg	atcatcacia	atcctttaa	acaattacct	2220
aaaagaaacc	aaaaatcctg	ccttctttgt	gggggagggg	ggagagaggg	gaaggaaatg	2280
gaacaagtgt	tgtttgtgtt	agcatgtggg	tgatgtaaac	ttcaaattgg	gagatgttcc	2340
gacccc						2346

<210> 660

<211> 2512

<212> DNA

<213> Homo sapiens

<400> 660

caatgcactg	acggatatga	gtgggatcct	gtgagacagc	aatgcaaaga	tattgatgaa	60
tgtgacattg	tcccagacgc	ttgtaaagg	ggaatgaagt	gtgtcaacca	ctatggagga	120
tacctctgcc	ttccgaaaac	agcccagatt	attgtcaata	atgaacagcc	tcagcaggaa	180
acacaaccag	cagaaggaac	ctcaggggca	accaccgggg	ttgtagctgc	cagcagcatg	240
gcaaccagtg	gagtgtttgc	cggggggtgt	tttgtggcca	gtgctgctgc	agtcgcaggc	300
cctgaaatgc	agactggccg	aaataacttt	gtcatccggc	ggaaccacgc	tgaccctcag	360
cgcattccct	ccaacccttc	ccaccgtatc	cagtgtgcag	caggctacga	gcaaagtga	420
cacaacgtgt	gccaagacat	agacgagtgc	actgcaggga	cgcacaactg	tagagcagac	480
caagtgtgca	tcaattttacg	gggatccttt	gcatgtcagt	gccctcctgg	atatcagaag	540
cgaggggagc	agtgcgtaga	catagatgaa	tgtaccatcc	ctccatattg	ccaccaaaga	600
tgcgtgaata	caccaggctc	atTTTtattgc	cagtgcagtc	ctgggtttca	attggcagca	660
aacaactata	cctgcgtaga	tataaatgaa	tgtgatgcca	gcaatcaatg	tgctcagcag	720
tgctacaaca	ttcttggttc	attcatctgt	cagtgcacac	aaggatatga	gctaagcagt	780
gacaggctca	actgtgaaga	cattgatgaa	tgcaaacct	caagctacct	gtgtcaatat	840
caatgtgtca	atgaacctgg	gaaattctca	tgtatgtgcc	cccagggata	ccaagtgggtg	900
agaagtagaa	catgtcaaga	tataaatgag	tgtgagacca	caaataaatg	ccgggaggat	960
gaaatgtgtt	ggaattatca	tggcggcttc	cgttgtttatc	cacgaaatcc	ttgtcaagat	1020
ccctacattc	taacaccaga	gaaccgatgt	gtttgcccag	tctcaaattgc	catgtgccga	1080
gaactgcccc	agtcaatagt	ctacaaatac	atgagcatcc	gatctgatag	gtctgtgcca	1140
tcagacatct	tccagatata	ggccacaact	atTTtatgcca	acaccatcaa	tacttttcgg	1200
attaaatctg	gaaatgaaaa	tggagagttc	tacctacgac	aaacaagtcc	tgtaagtgc	1260
atgcttgtgc	tcgtgaagtc	attatcagga	ccaagagaac	atatcgtgga	cctggagatg	1320
ctgacagtca	gcagtatagg	gaccttccgc	acaagctctg	tgTTaagatt	gacaataata	1380
gtggggccat	tttcatTTta	gtcttttcta	agagtcaacc	acaggcattt	aagtcagcca	1440
agaatatttg	ttaccttaaa	gcactatttt	atTTtatagat	atatctagtg	catctacatc	1500
tctatactgt	acactcacc	ataacaaaca	attacaccat	ggtataaagt	gggcatttaa	1560
tatgtaaaga	ttcaaagttt	gtctttatta	ctatatgtaa	attagacatt	aatccactaa	1620

actggtcttc	ttcaagagag	ctaagtatac	actatctggt	gaaacttgga	ttctttccta	1680
taaaagtggg	accaagcaat	gatgatcttc	tgtggtgctt	aaggaaactt	actagagctc	1740
cactaacagt	ctcataagga	ggcagccatc	ataaccattg	aatagcatgc	aagggtaaga	1800
atgagttttt	aactgctttg	taagaaaatg	gaaaagggtca	ataaagatat	atttcttttag	1860
aaaatgggga	tctgccatat	ttgtgttggt	ttttattttc	atatccagcc	taaagggtgg	1920
tgtttattat	atagtaataa	atcattgctg	tacaacatgc	tggtttctgt	agggtatttt	1980
taattttgtc	agaaatttta	gattgtgaat	attttgtaaa	aaacagtaag	caaaattttc	2040
cagaattccc	aaaatgaacc	agataccccc	tagaaaatta	tactattgag	aaatctatgg	2100
ggaggatatg	agaaaataaa	ttccttctaa	accacattgg	aactgacctg	aagaagcaaa	2160
ctcgaaaaat	ataataacat	ccctgaattc	aggcattcac	aagatgcaga	acaaaaatgga	2220
taaaagggtat	ttcactggag	aagttttaat	ttctaagtaa	aattttaaatc	ctaacacttc	2280
actaattttat	aactaaaatt	tctcatcttc	gtacttgatg	ctcacagagg	aagaaaaatga	2340
tgatggtttt	tattcctggc	atccagagtg	acagtgaact	taagcaaatt	accctcctac	2400
ccaattctat	ggaatathtt	atacgtctcc	ttgtttaaaa	tctgactgct	ttactttgat	2460
gtatcatatt	tttaataaaa	aataaatatt	cctttagaag	atcactctaa	aa	2512

<210> 661

<211> 2306

<212> DNA

<213> Homo sapiens

<400> 661

ggcggggagc	tgcacgcgcc	gtggctccgg	atctcttcgt	ctttgcagcg	tacgcccag	60
tcggtcagcg	ccggaggacc	tcagcagcca	tgtcgaagcc	ccatagtga	gccgggactg	120
ccttcattca	gaccagcag	ctgcacgcag	ccatggctga	cacattcctg	gagcacatgt	180
gccgcctgga	cattgattca	ccacccatca	cagcccgga	cactggcatc	atctgtacca	240
ttggcccagc	ttcccgatca	gtggagacgt	tgaaggagat	gattaagtct	ggaatgaatg	300
tggctcgtct	gaacttctct	catggaactc	atgagtacca	tgccggagacc	atcaagaatg	360
tgcgcacagc	cacggaaagc	tttgcttctg	acccctacct	ctaccggccc	gttgctgtgg	420
ctctagacac	taaaggacct	gagatccgaa	ctgggctcat	caagggcagc	ggcactgcag	480
agctggagct	gaagaaggga	gccactctca	aaatcacgct	ggataacgcc	tacatggaaa	540
agtgtgacga	gaacatcctg	tggctggact	acaagaacat	ctgcaagggtg	gtggaagtgg	600
gcagcaagat	ctacgtggat	gatgggctta	tttctctcca	ggtgaagcag	aaagggtgccg	660
acttctgggt	gacggagggtg	gaaaatgggtg	gtccttggg	cagcaagaag	ggtgtgaacc	720
ttcctggggc	tgctgtggac	ttgcctgctg	tgtcggagaa	ggacatccag	gatctgaagt	780
ttggggtcga	gcaggatgtt	gatatggtgt	ttgcgtcatt	catccgcaag	gcatctgatg	840
ttcatgaagt	taggaagggtc	ctgggagaga	agggaaagaa	catcaagatt	atcagcaaaa	900
tcgagaatca	tgagggggtt	cggagggttg	atgaaatcct	ggaggccagt	gatgggatca	960
tgggtggctcg	tgggtgacta	ggcattgaga	ttcctgcaga	gaagggtcttc	cttgctcaga	1020
agatgatgat	tggacggtgc	aaccgagctg	ggaagcctgt	catctgtgct	actcagatgc	1080
tggagagcat	gatcaagaag	ccccgcccc	ctcgggctga	aggcagtgat	gtggccaatg	1140
cagtcttgga	tggagccgac	tgcatcatgc	tgtctggaga	aacagccaaa	ggggactatc	1200
ctctggaggc	tgtgcgcatg	cagcacctga	ttgcccggtg	ggcagaggct	gccatctacc	1260
acttgcaatt	atttgaggaa	ctccgcccgc	tggcgcccat	taccagcgac	cccacagaag	1320
ccaccgccgt	gggtgccgtg	gaggcctcct	tcaagtgtctg	cagtggggcc	ataatcgtcc	1380
tcaccaagtc	tggcaggtct	gctcaccagg	tggccagata	ccgcccacgt	gcccccatca	1440
ttgctgtgac	ccggaatccc	cagacagctc	gtcaggccca	cctgtaccgt	ggcatcttcc	1500
ctgtgctgtg	caaggaccca	gtccaggagg	cctgggctga	ggacgtggac	ctccgggtga	1560
actttgccat	gaatgttggc	aaggcccag	gcttcttcaa	gaagggagat	gtgggtcattg	1620
tgctgaccgg	atggcgccct	ggctccggct	tcaccaaacac	catgcgtgtt	gttcctgtgc	1680
cgtgatggac	cccagagccc	ctcctccagc	ccctgtccca	cccccttccc	ccagcccac	1740
cattaggcca	gcaacgcttg	tagaactcac	tctgggctgt	aacgtggcac	tggtagggtg	1800
ggacaccagg	gaagaagatc	aacgcctcac	tgaacatgg	ctgtgtttgc	agcctgtct	1860
agtgggacag	cccagagcct	ggctgcccac	catgtggccc	cacccaatca	agggaagaag	1920

gaggaatgct	ggactggagg	cccctggagc	cagatggcaa	gaggggtgaca	gcttcctttc	1980
ctgtgtgtac	tctgtccagt	tccttttagaa	aaaatggatg	cccagaggac	tcccaaccct	2040
ggcttggggg	caagaaacag	ccagcaagag	ttagggggcct	tagggcactg	ggctgttggt	2100
ccattgaagc	cgactctggc	cctggccctt	acttgcttct	ctagctctct	aggcctctcc	2160
agtttgcacc	tgtccccacc	ctccactcag	ctgtcctgca	gcaaacactc	caccctccac	2220
cttccatttt	ccccactac	tgcagcacct	ccaggcctgt	tgctatagag	cctacctgta	2280
tgtcaataaa	caacagctga	agcacc				2306

<210> 662

<211> 2656

<212> DNA

<213> Homo sapiens

<400> 662

gaattcgcg	ccgcagagtc	cccggggccaa	gatggctgcy	cggtgctcca	cacgctgggt	60
gctgggtggt	gtgggggaccc	cgcggtgcc	ggctatatcg	ggtagagggg	cccggccgcc	120
caggggaggc	gtggtggggg	catggctgag	ccgcaagctg	agcgtccccg	cctttgctgc	180
ttccctgacc	tcttgcgggc	cccagagcgt	gctgacattg	agacctgggt	tcagccttac	240
aggaacaaaa	cataaccctt	tcatttgtag	tgctccttc	cacacgagtg	cccctttggc	300
caaagaagat	tattatcaga	tattaggagt	gcctcgaaat	gccagccaga	aagagatcaa	360
gaaagcctat	tatcagcttg	ccaagaagta	tcacctgac	acaaataagg	atgatcccaa	420
agccaaggag	aagttctccc	agctggcaga	agcctatgag	gttttgagtg	atgagggtgaa	480
gaggaagcag	tacgatgcct	acggctctgc	aggcttcgat	cctggggcca	gcggctccca	540
gcatagctac	tgggaaggag	gccccactgt	ggaccccgag	gagctgttca	ggaagatctt	600
tggcgagttc	tcctcctctt	catttgagga	tttccagacc	gtgtttgatc	agcctcagga	660
atacttcatg	gagttgacat	tcaatcaagc	tgcaaagggg	gtcaacaagg	agttcacctg	720
gaacatcatg	gacacgtgtg	agcgtgcaa	cggcaagggg	aacgagcccg	gcaccaaggt	780
gcagcattgc	cactactgtg	gcggctccgg	catggaaacc	atcaacacag	gcccttttgt	840
gatgcgttcc	acgtgtagga	gatgtggtgg	ccgcggctcc	atcatcatat	cgccctgtgt	900
ggtctgcagg	ggagcaggac	aagccaagca	gaaaaagcga	gtgatgatcc	ctgtgcctgc	960
aggagtcgag	gatggccaga	ccgtgaggat	gcctgtggga	aaaagggaaa	ttttcattac	1020
gttcaggggtg	cagaaaagcc	ctgtgttccg	gagggacggc	gcagacatcc	actccgacct	1080
ctttattttct	atagctcagg	ctcttcttgg	gggaacagcc	agagcccagg	gcctgtacga	1140
gacgatcaac	gtgacgatcc	cccctggggc	tcagacagac	cagaagattc	ggatgggtgg	1200
gaaaggcatc	ccccggatta	acagctacgg	ctacggagac	cactacatcc	acatcaagat	1260
acgagttcca	aagaggctaa	cgagccggca	gcagagcctg	atcctgagct	acgccgagga	1320
cgagacagat	gtggagggga	cggtgaacgg	cgtaaccctc	accagctctg	gtggcagcac	1380
catggatagc	tccgcaggaa	gcaaggctag	gcgtgaggct	ggggaggacg	aggagggtat	1440
cctttccaaa	cttaagaaaa	tgtttacctc	atgatatccc	agccgaggaa	aaagatcca	1500
tggaaaactag	gccgggaagc	agcagcccct	ccaagggcca	gggcacctgg	gagacgggag	1560
gattccagaa	cagcagcact	gagctcccac	ccgcagagcc	tctggacggc	cttggcaaca	1620
gcaaaatcat	gggacaacac	ctctctccac	ggaaagggtca	cagtggacag	cccgggcagt	1680
aggatgcagc	cccagaggct	ggtggcagtt	tctgtccat	tggtaggtga	cggccccctg	1740
gtcagcagag	gagagggttag	atcttgtagg	ctaaaactct	aatttggaat	tgaatattgt	1800
ggatatctta	gttaaaggcc	atgcttacag	cttagaaaatg	aagccttaag	ctgcatcaag	1860
ttacgaagtg	attaatttcc	ttctcagcaa	acctccggga	ggttccagaa	tgagttcttc	1920
ctgacaggtt	gtcttctactg	ggagcgtggg	gccccaggc	cccaccagca	ccgtcctccc	1980
ctaatagagg	gccttgccga	ggcatcagct	gctctgtctca	gttagttttt	attcccgggg	2040
taccaagcag	ctgcacagtc	ggtgcctggg	aagcacgtta	aaggcccaga	gagatcctgg	2100
gggttcttgc	ctgaccgtgt	gggtggtgat	ccttgtcagg	atgtacagtc	cttgcctcca	2160
ccccatccgg	gatggccggc	tgtccctgac	tattgagtcc	tggtgttgta	agccaggcat	2220
ggagggctcc	tgcccttctg	ctgagccaca	gccattgca	gcactgtgct	ggccagactt	2280
cagctgcctt	gggaactgaa	gccttgccac	tggtgctagt	caggggcttg	gttctccac	2340
ttacactggt	gacatctatt	ttctgaagtg	tggttaaatt	attcagtgct	aatcattggt	2400

ttttcctttg	taaatgttga	ttcagaaaag	gaaagcacag	gctaagcagt	tgaaggttcc	2460
ccaccattca	gtgagagcag	aacccccatt	ccccagcctc	tgctggtagc	atgtcgcagt	2520
ttccatgtgt	ttcaggatct	tcgggctgtc	gtagacagg	ttaatgaaga	acacttctca	2580
acagtttcct	ttttgttttc	ctttataatt	cactaaaata	aagcatctat	tagtgtctga	2640
aaaaaaaaaa	aaaaaa					2656

<210> 663

<211> 2217

<212> DNA

<213> Homo sapiens

<400> 663

gcggaccg	cgccgagg	gccacccg	acgcggcg	cacgctcc	cctgcgcag	60
ccggcccg	catggcgg	ccccgccg	ctccgcg	ctccgttt	gtctcggtc	120
cggttttta	cgccccg	aagaagt	gccctgtg	ggccccaa	cccaaagt	180
atcccttcc	gccccggg	agcgagcc	ccccggca	cgggggcc	cgcgcac	240
tgggcccgg	gggcgag	ccccgcgc	ccccgga	ctttccct	cctccac	300
cccttgctg	ggatggcg	gatgcag	gtgctctg	aggtgcct	ccgcgcgc	360
ctccccgat	cgaggaat	tttccctg	cgctctgg	ggaggaga	ttcccttc	420
cgccgcct	tccggagg	gagggagg	ctgaggcc	cataccgc	ccaccac	480
ccaggagaa	ggtgagc	attgattg	agatcgac	tctgtcct	ctgctgg	540
acatgacaa	gaatgat	ttcaaag	gggtgtca	tggatat	ccccacc	600
tggccaact	attcagtt	aagtccag	ccaagcct	agccgggg	acagcac	660
tgcctcctt	gaagtccc	tccagctc	agcctctg	ccaggttc	gctccgg	720
agagccag	acagttcc	gttcagcc	agccccag	caagcctc	gtccaact	780
atgtccag	ccagacc	cctgtgtc	tggtaaac	ccagcccc	gggcccc	840
cctcatct	ggctccag	cctaagtt	ctccagt	tcctaagt	actcctg	900
cttccaagt	cagtcctg	gccccagg	gatctggg	acaaccaa	caaaaatt	960
ggcacccca	agctcttt	gctggcac	gctccctc	acctccc	ttcacct	1020
cccagcag	ggagaag	cgagtgc	agaagcag	ccccgtg	ccaccgg	1080
agaaccaaa	ccagggtg	tccccggg	ccccagg	cctgactc	aaggagg	1140
aggagctg	gcagctga	cagcagct	tgcaggac	ggagcatc	cagaggca	1200
atgtggct	caacgaac	tgcggccg	gccatcaa	cctggccc	gcgcagcc	1260
ccgtccgc	tctagggc	ctgttcca	tcgcctgt	cacctgcc	cagtgtgc	1320
agcagctc	gggccagc	ttctacag	tggagggg	gccgtact	gagggctg	1380
acactgac	cctggaga	tgtaacac	gcggggag	catcactg	cgcatgtg	1440
gggccacgg	caaggcct	caccgcac	gcttcacct	tgtggctg	gcccgc	1500
tggagggc	ctccttc	gtggacc	ccaaccgg	ccactgtg	cccga	1560
acaagcag	cgccccg	tgtccgtc	gctctgag	catcatgc	gagcctg	1620
gagatgag	tgtgcgag	gtcgccct	acaagaac	ccacatga	tgttaca	1680
gtgaggact	cgggaa	ctgtcgat	aggcagat	caatggct	ttcccc	1740
acggtcac	gctctgtc	aagtgcc	ctgctag	ccagacct	gtgagg	1800
gccctctc	gaccgcag	catgcccc	tgtggacc	ccacactg	accacct	1860
cccacctc	ttattgtt	gatgtct	ccctccat	tccaaccc	ccctagc	1920
ccagggtc	tgaccag	cccaacat	tctaggg	caggatcc	gccctgg	1980
ctggtcct	cccatcct	agggattg	caccgtct	cagacac	acctgag	2040
ggcaccagg	ttagtgtg	tgctttc	gctgcacc	cgccctcg	cggcccc	2100
agcagcctt	gtactctg	tgcggagg	tgggagac	tccaggac	tcccac	2160
ccccatgct	ccaagttg	gctatag	caaataaaa	aaaacctt	tttccag	2217

<210> 664

<211> 8906

<212> DNA

<213> Homo sapiens

<400> 664

gaggcggcca	aggacctggc	cgacatcgcg	gccttcttcc	gatccggggtt	tcgaaaaaac	60
gatgaaatga	aagctatgga	tgttttacca	atthttgaagg	aaaaagttgc	atacctttca	120
ggtgggagag	ataaacgtgg	aggtcccatt	ttaacgtttc	cggcccgcag	caatcatgac	180
agaatacgac	aggaggatct	caggagactc	atttctatc	tagcctgtat	tcccagcgag	240
gaggtctgca	agcgtggctt	cacggtgatc	gtggacatgc	gtgggtccaa	gtgggactcc	300
atcaagcccc	ttctgaagat	cctgcaggag	tccttcccct	gctgcatcca	tgtggccctg	360
atcatcaagc	cagacaactt	ctggcagaaa	cagaggacta	atthttggcag	ttctaaatth	420
gaatthgaga	caaataatgg	ctctthtagaa	ggcctthacca	aagtagttga	tccttctcag	480
ctaactcctg	agthttgatgg	ctgcctggaa	tacaaccacg	aagaatggat	tgaaatcaga	540
gttgctthttg	aagactacat	tagcaatgcc	acccacatgc	tgtctcggct	ggaggaactt	600
caggacatcc	tagctaagaa	ggagctgcct	caggatthtag	agggggctcg	gaatatgatc	660
gaggaacatt	ctcagctgaa	gaagaagggtg	attaaggccc	ccatcgagga	cctggattht	720
gagggacaga	agctgcttca	gaggatacag	agcagtgaag	gctthcccaa	aaagaactca	780
ggctcaggca	atgcggacct	gcagaacctc	ttgcccaagg	tgtccaccat	gctggaccgg	840
ctgcactcga	cacggcagca	tctgcaccag	atgtggcatg	tgaggaagct	gaagctggac	900
cagtgccttc	agctgaggct	gtthgaacag	gatgctgaga	agatgtthga	ctggatcaca	960
cacaacaaag	gcctgtthct	aaacagctac	acagagatth	ggaccagcca	ccctcatgcc	1020
atggagcttc	agacgcagca	caatcactth	gccatgaact	gtatgaacgt	gtatgtaaat	1080
ataaacgcga	tcagtgcgg	ggccaatcgt	ctgggtggagt	ctggccacta	tgctctcgag	1140
cagatcaggc	agatcgcgag	tcagctggag	caggagtggg	aggcgtthtc	ggcagccctg	1200
gatgagcgga	gcaccttgct	ggacatgtcc	tcctthttcc	accagaaggc	cgaaaagtat	1260
atgagcaacg	tggattcatg	gtgtaaagct	tgcggtgagg	tagaccttcc	ctcagagctg	1320
caggacctag	aagatgccat	tcataccac	cagggaatat	atgaacatat	cactcttgct	1380
tattctgagg	tcagccaaga	tgggaagtgc	ctccttgaca	agctccagcg	gcccttgact	1440
cccgccagct	ccgattccct	gacagcctct	gccaactact	ccaaggccgt	gcaccatgtc	1500
ctggatgtca	tcacagaggt	gctgcaccac	cagcggcacg	tgagaacaat	ctggcaacac	1560
cgcaaggctc	ggctgcatca	gaggctgcag	ctgtgtgtth	tccagcagga	agttcagcag	1620
gtgctagact	ggatcgagaa	ccacggagaa	gcatttctga	gcaaacatac	aggtgtgggg	1680
aaatctcttc	atcggggccag	agcattgcag	aaacgtcatg	aagatthtga	agaagtggca	1740
cagaacacat	acaccaatgc	ggataaatta	ctggaagcag	cagaacagct	ggctcagact	1800
ggggaatgtg	accccgaa	gatttatcag	gctgcccac	agctggaaga	ccggattcaa	1860
gatttctgttc	ggcgtgttga	gcagcgaaag	atcctactgg	acatgtcagt	gtcctthtcac	1920
acccatgtga	aagagctgtg	gacgtggctg	gaggagctgc	agaaggagct	gctggacgac	1980
gtgtatgccg	agtcgggtga	ggcgtgcag	gacctcatca	agcgtthtgg	ccagcagcag	2040
cagaccaccc	tgcaggtgac	tgtcaacgtg	atcaaggaag	ggaggacct	catccagcag	2100
ctcagggact	ctgccatctc	cagtaacaag	acccccaca	acagctccat	caaccacatt	2160
gagacgggtg	tgcagcagct	ggacgaggcg	tgagcgaga	tgaggagct	cttcacaggag	2220
cgcaagatca	agctggagct	cttctctgac	gtgcgcactc	tcgagaggga	cgccatcgac	2280
attatctcag	acctcgagtc	ttggaatgat	gagctthtct	agcaaataga	tgacttctgac	2340
acagaagatc	tcacgattgc	agagcagcgc	ctccagcacc	atgcagacaa	agccttgacc	2400
atgaacaact	tgactthtga	cgtcatccac	caagggaag	atcttctgca	gtatgtcaat	2460
gaggtccagg	cctctgggtg	ggagctgctg	tgtgatagag	atgtagacat	ggcaactcgg	2520
gtccaggacc	tgctggagtt	tcttcatgaa	aaacagcagg	aattggatth	agccgcagag	2580
cagcatcgga	aacacctgga	gcagtgcgtg	cagctgcgcc	acctgcaggc	agaagtgaag	2640
caggtgctgg	gttggtatcc	caacggagag	tccatgttaa	atgccggact	tatcacagcc	2700
agctcggtac	aagaggcaga	gcagctccag	cgagagcacg	agcagttcca	gcatgccatt	2760
gagaaaacac	atcagagcgc	gctgcagggt	cagcagaagg	cagaagccat	gctacaggcc	2820
aaccactacg	acatggacat	gatccgggac	tgcgccgaga	aggtggcgct	tcactggcaa	2880
cagctcatgc	tcaagatgga	agatcgcttc	aagctcgctc	acgcctctgt	cgctthtctac	2940
aaaacctcag	agcaggtctg	cagcgtcctc	gagagcctgg	aacaggagta	caagagagaa	3000
gaagactggt	gtggcggggc	ggataagctg	ggcccaaact	ctgagacgga	ccacgtgacg	3060
cccattgatca	gcaagcacct	ggagcagaag	gaggcattcc	tgaaggcttg	caccttgct	3120
cggaggaatg	cagacgtctt	cctgaaatac	ctgcacagga	acagcgtgaa	catgccagga	3180

atggtgacgc	acatcaaagc	tcctgaacag	caagtgaaaa	atatcttgaa	tgaactcttc	3240
caacggggaga	acaggggtatt	gcattactgg	accatgagga	agagacggct	ggaccagtgt	3300
cagcagtagc	tgggtctttga	gaggagtgcc	aagcaggctt	tggaatggat	ccatgacaat	3360
ggcgagtctc	acctttccac	acacacctcc	acgggctcca	gtatacagca	cacccaggag	3420
ctcctgaaag	agcacgagga	gttcagata	actgcaaagc	aaaccaaaga	gagagtgaag	3480
ctattgatac	agctggctga	tggcttttgt	gaaaaagggc	atgcccattg	ggcagagata	3540
aaaaaatgtg	ttactgctgt	ggataagagg	tacagagatt	tctctctgcg	gatggagaag	3600
tacaggacct	ctttggaaaa	agccctgggg	atttcttcag	attccaacaa	atcgagtaaa	3660
agtctccagc	tagatatcat	tccagccagt	atccctgggt	cagagggtgaa	acttcgagat	3720
gctgctcatg	aacttaatat	agagaagcgg	aaatctgccc	gcaggaaaaga	gttcataatg	3780
gctgagctca	ttcaaaactga	aaaggcttat	gtaagagacc	tccgggaatg	tatggatacg	3840
tacctgtggg	aatgaccag	tggcgtggaa	gagattccac	ctggcattgt	aaacaaagaa	3900
ctcatcatct	tcggaacat	gcaagaaatc	tacgaatttc	ataataacat	attcctaaag	3960
gagctggaaa	aatatgaaca	gttgccagag	gatgttgagc	attgttttgt	tacttgggca	4020
gacaagtctc	agatgtatgt	cacatattgc	aaaaataagc	ctgattctac	tcagctgata	4080
ttggaacatg	cagggtccta	ttttgacgag	atacagcagc	gacatggatt	agccaattcc	4140
atttcttcct	accttattaa	accagttcag	cgaataacga	aatatcagct	ccttttataa	4200
gagctgctga	cgtgctgtga	ggaaggaaa	ggagagattt	aagatggcct	ggagggtgatg	4260
ctcagcgtgc	cgaagcgagc	caatgacgcc	atgcacctca	gcatgctgga	aggggttgat	4320
gaaaacattg	agtctcaggg	agaactcatc	ctacaggaat	ccttccaagt	gtgggaccca	4380
aaaaacctta	ttcgaaaggg	tcgagaacgg	catctcttcc	tttttgaaat	gtccttagta	4440
tttagtaaa	aagtgaaga	ttccagtggg	agaagcaagt	acctttataa	aagcaaatg	4500
tttacctcag	agttgggtgt	cacagaacat	gttgaaggag	accttgcaa	atttgactg	4560
tgggtgggga	gaacaccaac	ttcagataat	aaaattgtcc	ttaaggcttc	cagcatagag	4620
aacaagcagg	actggataaa	gcataatccg	gaagtcatcc	aggagcggac	gatccacctg	4680
aaggaggccc	tgaaggagcc	cattcacatc	cctaagaccg	ctcccggcac	aagacagaag	4740
ggaaggaggg	atggagagga	tctggacagc	caaggagacg	gcagcagcca	gcctgatacg	4800
atttccatcg	cctcacggac	gtctcagaac	acgctggaca	gcgataagct	ctctgggtggc	4860
tgtgagctga	cagtgggtgat	ccatgacttc	accgcttgca	acagcaacga	gctgaccatc	4920
cgacggggcc	agaccgtgga	agttctggag	cggccgcatg	acaagcctga	ctgggtgtctg	4980
gtgcggacca	ctgaccgctc	cccagcggca	gaaggcctgg	tcccctgtgg	ttcactgtgc	5040
atcgcccact	ccagaagtag	catggaaatg	gagggcatct	tcaaccacaa	agactcgctc	5100
tccgtctcca	gcaatgacgc	cagtccaccc	gcatccgtgg	cctccctcca	gccccacatg	5160
atcgggggccc	agagctcgcc	gggcccacag	cggccgggca	acaccctgcg	caagtggctc	5220
accagccccg	tgcggcggtc	cagcagcggc	aaggccgacg	ggcacgtgaa	gaagctggcg	5280
cacaagcaca	agaagagccg	cgaggtccgc	aagagcgccg	acgcccggctc	gcagaaggac	5340
tccgacgaca	gtgcggccac	cccgcaggag	gagagcgtcg	aggagagagg	ccggaacgag	5400
ggcctgagca	gcggtactct	ctccaaatcc	tcctctcgg	ggatgcagag	ctgtggagaa	5460
gaggaaggcg	aggagggggc	cgacgccgtg	cccctgccgc	cacccatggc	catccagcag	5520
cacagcctcc	tccagccaga	ctcacaggat	gacaaggcct	cttctcggtt	attagtccgc	5580
cccaccagct	ccgaaacacc	gagtgcagcc	gagctcgtca	gtgcaattga	ggaactcgtg	5640
aaaagcaaga	tggcactgga	ggatcgcccc	agctcactcc	ttgttgacca	gggagatagt	5700
agcagccctt	ccttcaaccc	ttcggataat	tcccttctct	cttctctctc	gcccattgat	5760
gagatggaag	aaaggaaatc	cagctcttta	aagagaagac	actacgtttt	gcaagaacta	5820
gtggagacag	agcgtgacta	tgtgcgggac	cttggctatg	tggttgaggg	ctacatggca	5880
cttatgaaag	aagatgggtg	tcctgatgac	atgaaaggaa	aagacaaaat	tgtgttcggc	5940
aacatccatc	agatttacga	ctggcacaga	gacttttttt	taggagagtt	agagaagtgc	6000
cttgaagatc	cagaaaaact	aggatccctt	tttgttaaac	acgagagaag	gttgccacatg	6060
tacatagctt	attgtcaaaa	taaaccaaa	tctgagcaca	ttgtctcaga	atacattgat	6120
accttttttg	aggacttaaa	gcagcgtctt	ggccacagg	tacagctcac	agatctgttg	6180
atcaaaaccg	tgagagaaat	catgaagtat	cagctgttac	tgaaggactt	cctcaagtat	6240
tccaaaaaagg	ccagcctgga	tacatcagaa	ttagagagag	ctgtggaagt	catgtgcata	6300
gtacccaggc	ggtgcaacga	catgatgaac	gtggggcggc	tgcaaggatt	cgacgggaaa	6360
atcggtgccc	agggtaaaat	gctcttgacg	gacacattct	tggtcacaga	ccaagatgca	6420

ggactttctgc	ctcgctgcag	agagagggcg	atcttcctct	ttgagcagat	cgtcatattc	6480
agcgaaccac	ttgataaaaa	gaagggcttc	tccatgccgg	gattcctgtt	taagaacagt	6540
atcaaggtga	gttgcccttg	cctggaggaa	aatgtggaaa	atgatccctg	taaatttgct	6600
ctgacatcga	ggacgggtga	cgtggtagag	accttcattt	tgcatcctc	tagtccaagt	6660
gtccggcaaa	cttggatcca	tgaatcaac	caaatttttag	aaaaccagcg	caatttttta	6720
aatgccttga	catcgccaat	cgagtaccag	aggaaccaca	gcggggggcg	cgggcgggcg	6780
ggcagcgggg	cagcgggggg	ggtgggggca	gcggcgggcg	cgggggcccc	agtggcgggc	6840
gcggccacag	tggcgggccc	agcagctgcg	gcggcgcccc	cagcacgagc	aggagcgggc	6900
cctcccggat	cccccagcct	gtccgacacc	acccccccgt	gctggtctcc	tctgcagcct	6960
cgagccaggc	agaggcagac	aagatgtcag	agtgaagca	gcagcagtag	caacatctcc	7020
accatgttgg	tgacacacga	ttacacggca	gtgaaggagg	atgagatcaa	cgtctaccaa	7080
ggagaggtcg	ttcaaattct	ggccagcaac	cagcagaaca	tgtttctggt	gttccgagcc	7140
gccactgacc	agtgcctcgc	agctgagggc	tggattccag	gctttgtcct	ggggccacacc	7200
agtgcagtca	tcgtggagaa	cccggacggg	actctcaaga	agtcaacatc	ttggcacaca	7260
gcactccgtt	taaggaaaaa	atctgagaaa	aaagataaag	acggcaaaaag	ggaaggcaag	7320
ttagagaacg	gttatcgga	gtcacgggaa	ggactcagca	acaaggatct	tgtgaagctt	7380
ctcaatccca	actacattta	tgacgttccc	ccagaattcg	tcattccatt	gagtggagtc	7440
acgtgtgaga	caggggagac	cgttggttct	agatgtcgag	tctgtggccg	cccaaagcc	7500
tcaattacct	ggaagggccc	tgaacacaac	accttgaaca	acgatggtca	ctacagcatc	7560
tcctacagtg	acctgggaga	ggccacgctg	aagattgtgg	gcgtgaccac	ggaagatgac	7620
ggcatctaca	cgtgcctcgc	tgtcaatgac	atgggttcag	cctcatcctc	ggccagcctg	7680
agggctctag	gtccagggat	ggatgggatc	atgggtgacct	ggaaagacaa	ctttgactcc	7740
ttctacagtg	aagtggctga	gcttggcagg	ggcagattct	ctgtcgttaa	gaaatgtgat	7800
cagaaaggaa	ccaagcgagc	agtggccact	aagtttgtga	acaagaagtt	gatgaagcgc	7860
gaccaggtca	cccatgagct	tggcatcctg	cagagcctcc	agcaccctct	gcttgtcggc	7920
ctcctcgaca	cctttgagac	ccccaccagc	tacatcctgg	tcttagaaat	ggctgaccag	7980
ggtcgccctc	tggactgcgt	ggtgcgatgg	ggaagcctca	ctgaagggaa	gatcagggcg	8040
cacctggggg	aggttctgga	agctgtccgg	tacctgcaca	actgcaggat	agcacacctg	8100
gacctaaagc	ctgagaatat	cctggtggat	gagagttag	ccaagccaac	catcaaactg	8160
gctgactttg	gagatgctgt	tcagctcaac	acgacctact	acatccacca	gttactgggg	8220
aacctgaat	tcgcagcccc	tgaatcatc	ctcggaacc	ctgtctccct	gacctcggat	8280
acgtggagtg	ttggagtgc	cacatacgta	cttcttagtg	gcgtgtcccc	cttcctggat	8340
gacagtgtgg	aagagacctg	cctgaacatt	tgccgcttag	actttagctt	cccagatgac	8400
tactttaaag	gagtgaacca	gaaggccaag	gagttcgtgt	gcttcctcct	gcaggaggac	8460
cccggcaagc	gtccctcggc	tgcgctggcc	ctccaggagc	agtggctgca	ggccggcaac	8520
ggcagaagca	cgggcgtcct	cgacacgtcc	agactgactt	ccttcattga	gcggcgcaaa	8580
caccagaatg	atgttcgacc	tatccgtagc	attaaaaact	ttctgcagag	caggcttctg	8640
cttagagttt	gacctatcca	gaagttcttt	ctcattctct	ttcacctgcc	aatcagctgt	8700
taatctgaat	tttcaagaga	aaacaagcaa	acataactga	tcagctgccg	gtatgttcat	8760
cgtgtgaaat	tgcattccaa	gtgagctgtg	ctcagcagtg	cttggacaca	gagctgcaag	8820
ctgcgctggg	gtggaggacc	gtcacttaca	ctctgccaa	gacggaggtc	gcattgtctg	8880
atcacagtat	tttttacgga	tttctg				8906

<210> 665

<211> 1204

<212> DNA

<213> Homo sapiens

<400> 665

tcggcgggcg	tggatcggc	ggcagctgtg	aggggggttc	gggaagatgg	tgctgatcaa	60
ggaattccgt	gtggttttgc	catgttctgt	tcaggagtat	caggttgggc	agctttactc	120
tgttgagaa	gctagtaaga	atgagactgg	tgggtggagaa	ggaattgaag	tcttaaagaa	180
tgaaccttat	gagaaggatg	gagaaaagg	acagtatacg	cacaaaattt	atcacctaaa	240

```

gagcaaaagtg cctgcattcg tgaggatgat tgctcccagag ggctccttgg tgtttcatga 300
gaaaagcctgg aatgcgtacc cctactgtag aacaattgta acgaatgaat atatgaaaga 360
tgattttcttc attaaaaatcg aaacatggca caaaccagac ttgggaacat tagaaaatgt 420
acatgggttta gatccaaaca catggaaaac tgttgaaatt gtccatatag atattgcaga 480
tagaagtcaa gttgaaccag cagactacaa agctgatgaa gaccagcat tattccagtc 540
agtcaagacc aagagaggcc ctttgggacc caactggaag aaggagctgg caaacagccc 600
tgactgtccc cagatgtgtg cctataagct ggtgaccatc aaattcaagt ggtggggact 660
gcaaagcaaa gtagaaaact tcattcaaaa gcaagaaaaa cggatattta caaacttcca 720
tcgccagctt ttttgttga ttgacaagtg gatcgatctc acgatggaag acattaggag 780
aatggaagac gagactcaga aagaactaga aacaatgcgt aagaggggtt ccgttcgagg 840
cacgtcggct gctgatgtct agatgagtc cctgtagggg cagagacaat gtcaaactgt 900
ttacgtaatc aagggtcaagt gaggggaaca agcgcagcca gtgatgagtg aacaacaatc 960
tgaccagtat cttgcagtgt tgacgtttcc cagatgtgtg cttgtgatga tacacacaca 1020
tgcacagggt ctcaaccacg tgtgtatata tgtatgtgtg catatgtctg tagctgtata 1080
taaagcgcgt gtagagctac agatccagat acacacactt gtgtatatat gtacatacac 1140
acatactgaa gggattagta caatttctcc aaagtactgt acctatcttc agcaagaatg 1200
caaa 1204

```

<210> 666

<211> 892

<212> PRT

<213> Homo sapiens

<400> 666

```

Met Asp His Tyr Asp Ser Gln Gln Thr Asn Asp Tyr Met Gln Pro Glu
 1          5          10          15
Glu Asp Trp Asp Arg Asp Leu Leu Leu Asp Pro Ala Trp Glu Lys Gln
          20          25          30
Gln Arg Lys Thr Phe Thr Ala Trp Cys Asn Ser His Leu Arg Lys Ala
          35          40          45
Gly Thr Gln Ile Glu Asn Ile Glu Glu Asp Phe Arg Asp Gly Leu Lys
          50          55          60
Leu Met Leu Leu Leu Glu Val Ile Ser Gly Glu Arg Leu Ala Lys Pro
65          70          75          80
Glu Arg Gly Lys Met Arg Val His Lys Ile Ser Asn Val Asn Lys Ala
          85          90          95
Leu Asp Phe Ile Ala Ser Lys Gly Val Lys Leu Val Ser Ile Gly Ala
          100          105          110
Glu Glu Ile Val Asp Gly Asn Val Lys Met Thr Leu Gly Met Ile Trp
          115          120          125
Thr Ile Ile Leu Arg Phe Ala Ile Gln Asp Ile Ser Val Glu Glu Thr
          130          135          140
Ser Ala Lys Glu Gly Leu Leu Leu Trp Cys Gln Arg Lys Thr Ala Pro
145          150          155          160
Tyr Lys Asn Val Asn Ile Gln Asn Phe His Ile Ser Trp Lys Asp Gly
          165          170          175
Leu Gly Phe Cys Ala Leu Ile His Arg His Arg Pro Glu Leu Ile Asp
          180          185          190
Tyr Gly Lys Leu Arg Lys Asp Asp Pro Leu Thr Asn Leu Asn Thr Ala
          195          200          205
Phe Asp Val Ala Glu Lys Tyr Leu Asp Ile Pro Lys Met Leu Asp Ala
          210          215          220
Glu Asp Ile Val Gly Thr Ala Arg Pro Asp Glu Lys Ala Ile Met Thr
225          230          235          240

```

Tyr Val Ser Ser Phe Tyr His Ala Phe Ser Gly Ala Gln Lys Ala Glu
 245 250 255
 Thr Ala Ala Asn Arg Ile Cys Lys Val Leu Ala Val Asn Gln Glu Asn
 260 265 270
 Glu Gln Leu Met Glu Asp Tyr Glu Lys Leu Ala Ser Asp Leu Leu Glu
 275 280 285
 Trp Ile Arg Arg Thr Ile Pro Trp Leu Glu Asn Arg Val Pro Glu Asn
 290 295 300
 Thr Met His Ala Met Gln Gln Lys Leu Glu Asp Phe Arg Asp Tyr Arg
 305 310 315 320
 Arg Leu His Lys Pro Pro Lys Val Gln Glu Lys Cys Gln Leu Glu Ile
 325 330 335
 Asn Phe Asn Thr Leu Gln Thr Lys Leu Arg Leu Ser Asn Arg Pro Ala
 340 345 350
 Phe Met Pro Ser Glu Gly Arg Met Val Ser Asp Ile Asn Asn Ala Trp
 355 360 365
 Gly Cys Leu Glu Gln Val Glu Lys Gly Tyr Glu Glu Trp Leu Leu Asn
 370 375 380
 Glu Ile Arg Arg Leu Glu Arg Leu Asp His Leu Ala Glu Lys Phe Arg
 385 390 395 400
 Gln Lys Ala Ser Ile His Glu Ala Trp Thr Asp Gly Lys Glu Ala Met
 405 410 415
 Leu Arg Gln Lys Asp Tyr Glu Thr Ala Thr Leu Ser Glu Ile Lys Ala
 420 425 430
 Leu Leu Lys Lys His Glu Ala Phe Glu Ser Asp Leu Ala Ala His Gln
 435 440 445
 Asp Arg Val Glu Gln Ile Ala Ala Ile Ala Gln Glu Leu Asn Glu Leu
 450 455 460
 Asp Tyr Tyr Asp Ser Pro Ser Val Asn Ala Arg Cys Gln Lys Ile Cys
 465 470 475 480
 Asp Gln Trp Asp Asn Leu Gly Ala Leu Thr Gln Lys Arg Arg Glu Ala
 485 490 495
 Leu Glu Arg Thr Glu Lys Leu Leu Glu Thr Ile Asp Gln Leu Tyr Leu
 500 505 510
 Glu Tyr Ala Lys Arg Ala Ala Pro Phe Asn Asn Trp Met Glu Gly Ala
 515 520 525
 Met Glu Asp Leu Gln Asp Thr Phe Ile Val His Thr Ile Glu Glu Ile
 530 535 540
 Gln Gly Leu Thr Thr Ala His Glu Gln Phe Lys Ala Thr Leu Pro Asp
 545 550 555 560
 Ala Asp Lys Glu Arg Leu Ala Ile Leu Gly Ile His Asn Glu Val Ser
 565 570 575
 Lys Ile Val Gln Thr Tyr His Val Asn Met Ala Gly Thr Asn Pro Tyr
 580 585 590
 Thr Thr Ile Thr Pro Gln Glu Ile Asn Gly Lys Trp Asp His Val Arg
 595 600 605
 Gln Leu Val Pro Arg Arg Asp Gln Ala Leu Thr Glu Glu His Ala Arg
 610 615 620
 Gln Gln His Asn Glu Arg Leu Arg Lys Gln Phe Gly Ala Gln Ala Asn
 625 630 635 640
 Val Ile Gly Pro Trp Ile Gln Thr Lys Met Glu Glu Ile Gly Arg Ile
 645 650 655
 Ser Ile Glu Met His Gly Thr Leu Glu Asp Gln Leu Ser His Leu Arg
 660 665 670

Gln Tyr Glu Lys Ser Ile Val Asn Tyr Lys Pro Lys Ile Asp Gln Leu
 675 680 685
 Glu Gly Asp His Gln Leu Ile Gln Glu Ala Leu Ile Phe Asp Asn Lys
 690 695 700
 His Thr Asn Tyr Thr Met Glu His Ile Arg Val Gly Trp Glu Gln Leu
 705 710 715 720
 Leu Thr Thr Ile Ala Arg Thr Ile Asn Glu Val Glu Asn Gln Ile Leu
 725 730 735
 Thr Arg Asp Ala Lys Gly Ile Ser Gln Glu Gln Met Asn Glu Phe Arg
 740 745 750
 Ala Ser Phe Asn His Phe Asp Arg Asp His Ser Gly Thr Leu Gly Pro
 755 760 765
 Glu Glu Phe Lys Ala Cys Leu Ile Ser Leu Gly Tyr Asp Ile Gly Asn
 770 775 780
 Asp Pro Gln Gly Glu Ala Glu Phe Ala Arg Ile Met Ser Ile Val Asp
 785 790 795 800
 Pro Asn Arg Leu Gly Val Val Thr Phe Gln Ala Phe Ile Asp Phe Met
 805 810 815
 Ser Arg Glu Thr Ala Asp Thr Asp Thr Ala Asp Gln Val Met Ala Ser
 820 825 830
 Phe Lys Ile Leu Ala Gly Asp Lys Asn Tyr Ile Thr Met Asp Glu Leu
 835 840 845
 Arg Arg Glu Leu Pro Pro Asp Gln Ala Glu Tyr Cys Ile Ala Arg Met
 850 855 860
 Ala Pro Tyr Thr Gly Pro Asp Ser Val Pro Gly Ala Leu Asp Tyr Met
 865 870 875 880
 Ser Phe Ser Thr Ala Leu Tyr Gly Glu Ser Asp Leu
 885 890

<210> 667

<211> 197

<212> PRT

<213> Homo sapiens

<400> 667

Met Met Phe Pro Gln Ser Arg His Ser Gly Ser Ser His Leu Pro Gln
 1 5 10 15
 Gln Leu Lys Phe Thr Thr Ser Asp Ser Cys Asp Arg Ile Lys Asp Glu
 20 25 30
 Phe Gln Leu Leu Gln Ala Gln Tyr His Ser Leu Lys Leu Glu Cys Asp
 35 40 45
 Lys Leu Ala Ser Glu Lys Ser Glu Met Gln Arg His Tyr Val Met Tyr
 50 55 60
 Tyr Glu Met Ser Tyr Gly Leu Asn Ile Glu Met His Lys Gln Ala Glu
 65 70 75 80
 Ile Val Lys Arg Leu Asn Gly Ile Cys Ala Gln Val Leu Pro Tyr Leu
 85 90 95
 Ser Gln Glu His Gln Gln Gln Val Leu Gly Ala Ile Glu Arg Ala Lys
 100 105 110
 Gln Val Thr Ala Pro Glu Leu Asn Ser Ile Ile Arg Gln Gln Leu Gln
 115 120 125
 Ala His Gln Leu Ser Gln Leu Gln Ala Leu Ala Leu Pro Leu Thr Pro
 130 135 140
 Leu Pro Val Gly Leu Gln Pro Pro Ser Leu Pro Ala Val Ser Ala Gly

145 150 155 160
 Thr Gly Leu Leu Ser Leu Ser Ala Leu Gly Ser Gln Ala His Leu Ser
 165 170 175
 Lys Glu Asp Lys Asn Gly His Asp Gly Asp Thr His Gln Glu Asp Asp
 180 185 190
 Gly Glu Lys Ser Asp
 195

<210> 668

<211> 739

<212> PRT

<213> Homo sapiens

<400> 668

Gly Asp Lys Glu Pro Thr Glu Thr Ile Gly Asp Leu Ser Ile Cys Leu
 1 5 10 15
 Asp Gly Leu Gln Leu Glu Ser Glu Val Val Thr Asn Gly Glu Thr Thr
 20 25 30
 Cys Ser Glu Ser Ala Ser Gln Asn Asp Asp Gly Ser Arg Ser Lys Asp
 35 40 45
 Glu Thr Arg Val Ser Thr Asn Gly Ser Asp Asp Pro Glu Asp Ala Gly
 50 55 60
 Ala Gly Glu Asn Arg Arg Val Ser Gly Asn Asn Ser Pro Ser Leu Ser
 65 70 75 80
 Asn Gly Gly Phe Lys Pro Ser Arg Pro Pro Arg Pro Ser Arg Pro Pro
 85 90 95
 Pro Pro Thr Pro Arg Arg Pro Ala Ser Val Asn Gly Ser Pro Ser Ala
 100 105 110
 Thr Ser Glu Ser Asp Gly Ser Ser Thr Gly Ser Leu Pro Pro Thr Asn
 115 120 125
 Thr Asn Thr Asn Thr Ser Glu Gly Ala Thr Ser Gly Leu Ile Ile Pro
 130 135 140
 Leu Thr Ile Ser Gly Gly Ser Gly Pro Arg Pro Leu Asn Pro Val Thr
 145 150 155 160
 Gln Ala Pro Leu Pro Pro Gly Trp Glu Gln Arg Val Asp Gln His Gly
 165 170 175
 Arg Val Tyr Tyr Val Asp His Val Glu Lys Arg Thr Thr Trp Asp Arg
 180 185 190
 Pro Glu Pro Leu Pro Pro Gly Trp Glu Arg Arg Val Asp Asn Met Gly
 195 200 205
 Arg Ile Tyr Tyr Val Asp His Phe Thr Arg Thr Thr Trp Gln Arg
 210 215 220
 Pro Thr Leu Glu Ser Val Arg Asn Tyr Glu Gln Trp Gln Leu Gln Arg
 225 230 235 240
 Ser Gln Leu Gln Gly Ala Met Gln Gln Phe Asn Gln Arg Phe Ile Tyr
 245 250 255
 Gly Asn Gln Asp Leu Phe Ala Thr Ser Gln Ser Lys Glu Phe Asp Pro
 260 265 270
 Leu Gly Pro Leu Pro Pro Gly Trp Glu Lys Arg Thr Asp Ser Asn Gly
 275 280 285
 Arg Val Tyr Phe Val Asn His Asn Thr Arg Ile Thr Gln Trp Glu Asp
 290 295 300
 Pro Arg Ser Gln Gly Gln Leu Asn Glu Lys Pro Leu Pro Glu Gly Trp
 305 310 315 320

Glu Met Arg Phe Thr Val Asp Gly Ile Pro Tyr Phe Val Asp His Asn
 325 330 335
 Arg Arg Thr Thr Thr Tyr Ile Asp Pro Arg Thr Gly Lys Ser Ala Leu
 340 345 350
 Asp Asn Gly Pro Gln Ile Ala Tyr Val Arg Asp Phe Lys Ala Lys Val
 355 360 365
 Gln Tyr Phe Arg Phe Trp Cys Gln Gln Leu Ala Met Pro Gln His Ile
 370 375 380
 Lys Ile Thr Val Thr Arg Lys Thr Leu Phe Glu Asp Ser Phe Gln Gln
 385 390 395 400
 Ile Met Ser Phe Ser Pro Gln Asp Leu Arg Arg Arg Leu Trp Val Ile
 405 410 415
 Phe Pro Gly Glu Glu Gly Leu Asp Tyr Gly Gly Val Ala Arg Glu Trp
 420 425 430
 Phe Phe Leu Leu Ser His Glu Val Leu Asn Pro Met Tyr Cys Leu Phe
 435 440 445
 Glu Tyr Ala Gly Lys Asp Asn Tyr Cys Leu Gln Ile Asn Pro Ala Ser
 450 455 460
 Tyr Ile Asn Pro Asp His Leu Lys Tyr Phe Arg Phe Ile Gly Arg Phe
 465 470 475 480
 Ile Ala Met Ala Leu Phe His Gly Lys Phe Ile Asp Thr Gly Phe Ser
 485 490 495
 Leu Pro Phe Tyr Lys Arg Ile Leu Asn Lys Pro Val Gly Leu Lys Asp
 500 505 510
 Leu Glu Ser Ile Asp Pro Glu Phe Tyr Asn Ser Leu Ile Trp Val Lys
 515 520 525
 Glu Asn Asn Ile Glu Glu Cys Asp Leu Glu Met Tyr Phe Ser Val Asp
 530 535 540
 Lys Glu Ile Leu Gly Glu Ile Lys Ser His Asp Leu Lys Pro Asn Gly
 545 550 555 560
 Gly Asn Ile Leu Val Thr Glu Glu Asn Lys Glu Glu Tyr Ile Arg Met
 565 570 575
 Val Ala Glu Trp Arg Leu Ser Arg Gly Val Glu Glu Gln Thr Gln Ala
 580 585 590
 Phe Phe Glu Gly Phe Asn Glu Ile Leu Pro Gln Gln Tyr Leu Gln Tyr
 595 600 605
 Phe Asp Ala Lys Glu Leu Glu Val Leu Leu Cys Gly Met Gln Glu Ile
 610 615 620
 Asp Leu Asn Asp Trp Gln Arg His Ala Ile Tyr Arg His Tyr Ala Arg
 625 630 635 640
 Thr Ser Lys Gln Ile Met Trp Phe Trp Gln Phe Val Lys Glu Ile Asp
 645 650 655
 Asn Glu Lys Arg Met Arg Leu Leu Gln Phe Val Thr Gly Thr Cys Arg
 660 665 670
 Leu Pro Val Gly Gly Phe Ala Asp Leu Met Gly Ser Asn Gly Pro Gln
 675 680 685
 Lys Phe Cys Ile Glu Lys Val Gly Lys Glu Asn Trp Leu Pro Arg Ser
 690 695 700
 His Thr Cys Phe Asn Arg Leu Asp Leu Pro Pro Tyr Lys Ser Tyr Glu
 705 710 715 720
 Gln Leu Lys Glu Lys Leu Leu Phe Ala Ile Glu Glu Thr Glu Gly Phe
 725 730 735
 Gly Gln Glu

<210> 669
 <211> 431
 <212> PRT
 <213> Homo sapiens

<400> 669

```

Gly Pro Pro Pro Thr Arg Ala Leu Pro Leu Pro Gln Ser Leu Pro Pro
 1          5          10          15
Asp Phe Arg Leu Glu Pro Thr Ala Pro Ala Leu Ser Pro Arg Ser Ser
 20          25          30
Phe Ala Ser Ser Ser Ala Ser Asp Ala Ser Lys Pro Ser Ser Pro Arg
 35          40          45
Gly Ser Leu Leu Leu Asp Gly Ala Gly Ala Gly Gly Ala Gly Gly Ser
 50          55          60
Arg Pro Cys Ser Asn Arg Thr Ser Gly Ile Ser Met Gly Tyr Asp Gln
 65          70          75          80
Arg His Gly Ser Pro Leu Pro Ala Gly Pro Cys Leu Phe Gly Pro Pro
 85          90          95
Leu Ala Gly Ala Pro Ala Gly Tyr Ser Pro Gly Gly Val Pro Ser Ala
 100         105         110
Tyr Pro Glu Leu His Ala Ala Leu Asp Arg Leu Tyr Ala Gln Arg Pro
 115        120        125
Ala Gly Phe Gly Cys Gln Glu Ser Arg His Ser Tyr Pro Pro Ala Leu
 130        135        140
Gly Ser Pro Gly Ala Leu Ala Gly Ala Arg Val Gly Ala Ala Gly Pro
 145        150        155        160
Leu Glu Arg Arg Gly Ala Gln Pro Gly Arg His Ser Val Thr Gly Tyr
 165        170        175
Gly Asp Cys Ala Val Gly Ala Arg Tyr Gln Asp Glu Leu Thr Ala Leu
 180        185        190
Leu Arg Leu Thr Val Gly Thr Gly Gly Arg Glu Ala Gly Ala Arg Gly
 195        200        205
Glu Pro Ser Gly Ile Glu Pro Ser Gly Leu Glu Glu Pro Pro Gly Pro
 210        215        220
Phe Val Pro Glu Ala Ala Arg Ala Arg Met Arg Glu Pro Glu Ala Arg
 225        230        235        240
Glu Asp Tyr Phe Gly Thr Cys Ile Lys Cys Asn Lys Gly Ile Tyr Gly
 245        250        255
Gln Ser Asn Ala Cys Gln Ala Leu Asp Ser Leu Tyr His Thr Gln Cys
 260        265        270
Phe Val Cys Cys Ser Cys Gly Arg Thr Leu Arg Cys Lys Ala Phe Tyr
 275        280        285
Ser Val Asn Gly Ser Val Tyr Cys Glu Glu Asp Tyr Leu Phe Ser Gly
 290        295        300
Phe Gln Glu Ala Ala Glu Lys Cys Cys Val Cys Gly His Leu Ile Leu
 305        310        315        320
Glu Lys Ile Leu Gln Ala Met Gly Lys Ser Tyr His Pro Gly Cys Phe
 325        330        335
Arg Cys Ile Val Cys Asn Lys Cys Leu Asp Gly Ile Pro Phe Thr Val
 340        345        350
Asp Phe Ser Asn Gln Val Tyr Cys Val Thr Asp Tyr His Lys Asn Tyr
 355        360        365
Ala Pro Lys Cys Ala Ala Cys Gly Gln Pro Ile Leu Pro Ser Glu Gly
 370        375        380

```

Cys Glu Asp Ile Val Arg Val Ile Ser Met Asp Arg Asp Tyr His Phe
 385 390 395 400
 Glu Cys Tyr His Cys Glu Asp Cys Arg Met Gln Leu Ser Asp Glu Glu
 405 410 415
 Gly Cys Cys Cys Phe Pro Leu Asp Gly His Leu Leu Cys His Gly
 420 425 430

<210> 670

<211> 900

<212> PRT

<213> Homo sapiens

<400> 670

Gly Pro Gly Ser Arg His His Arg Ala Arg Asp Arg Leu Ile His Phe
 1 5 10 15
 Gly Ala Val Ser Thr Asp Val Leu Gly Cys Ser Ala His Cys Ser Leu
 20 25 30
 Thr Gln Ser Pro Lys Met Asn Ile Gln Glu Gln Gly Phe Pro Leu Asp
 35 40 45
 Leu Gly Ala Ser Phe Thr Glu Asp Ala Pro Arg Pro Pro Val Pro Gly
 50 55 60
 Glu Glu Gly Glu Leu Val Ser Thr Asp Pro Arg Pro Ala Ser Tyr Ser
 65 70 75 80
 Phe Cys Ser Gly Lys Gly Val Gly Ile Lys Gly Glu Thr Ser Thr Ala
 85 90 95
 Thr Pro Arg Arg Ser Asp Leu Asp Leu Gly Tyr Glu Pro Glu Gly Ser
 100 105 110
 Ala Ser Pro Thr Pro Pro Tyr Leu Lys Trp Ala Glu Ser Leu His Ser
 115 120 125
 Leu Leu Asp Asp Gln Asp Gly Ile Ser Leu Phe Arg Thr Phe Leu Lys
 130 135 140
 Gln Glu Gly Cys Ala Asp Leu Leu Asp Phe Trp Phe Ala Cys Thr Gly
 145 150 155 160
 Phe Arg Lys Leu Glu Pro Cys Asp Ser Asn Glu Glu Lys Arg Leu Lys
 165 170 175
 Leu Ala Arg Ala Ile Tyr Arg Lys Tyr Ile Leu Asp Asn Asn Gly Ile
 180 185 190
 Val Ser Arg Gln Thr Lys Pro Ala Thr Lys Ser Phe Ile Lys Gly Cys
 195 200 205
 Ile Met Lys Gln Leu Ile Asp Pro Ala Met Phe Asp Gln Ala Gln Thr
 210 215 220
 Glu Ile Gln Ala Thr Met Glu Glu Asn Thr Tyr Pro Ser Phe Leu Lys
 225 230 235 240
 Ser Asp Ile Tyr Leu Glu Tyr Thr Arg Thr Gly Ser Glu Ser Pro Lys
 245 250 255
 Val Cys Ser Asp Gln Ser Ser Gly Ser Gly Thr Gly Lys Gly Ile Ser
 260 265 270
 Gly Tyr Leu Pro Thr Leu Asn Glu Asp Glu Glu Trp Lys Cys Asp Gln
 275 280 285
 Asp Met Asp Glu Asp Asp Gly Arg Asp Ala Ala Pro Pro Gly Arg Leu
 290 295 300
 Pro Gln Lys Leu Leu Leu Glu Thr Ala Ala Pro Arg Val Ser Ser Ser
 305 310 315 320
 Arg Arg Tyr Ser Glu Gly Arg Glu Phe Arg Tyr Gly Ser Trp Arg Glu

-225-

```

      755              760              765
Val Gln Glu Val Met Arg Arg Gly Arg Ala Cys Val Arg Pro Ala Cys
      770              775              780
Ala Pro Val Leu His Val Val Pro Ala Val Ser Asp Met Glu Leu Ser
      785              790              795              800
Glu Thr Glu Thr Arg Ser Gln Arg Lys Val Gly Gly Gly Ser Ala Gln
      805              810              815
Pro Cys Asp Ser Ile Val Val Ala Tyr Tyr Phe Cys Gly Glu Pro Ile
      820              825              830
Pro Tyr Arg Thr Leu Val Arg Gly Arg Ala Val Thr Leu Gly Gln Phe
      835              840              845
Lys Glu Leu Leu Thr Lys Lys Gly Ser Tyr Arg Tyr Tyr Phe Lys Lys
      850              855              860
Val Ser Asp Glu Phe Asp Cys Gly Val Val Phe Glu Glu Val Arg Glu
      865              870              875              880
Asp Glu Ala Val Leu Pro Val Phe Glu Glu Lys Ile Ile Gly Lys Val
      885              890              895
Glu Lys Val Asp
      900

```

<210> 671

<211> 591

<212> PRT

<213> Homo sapiens

<400> 671

```

Met Val Pro Val Ala Val Thr Ala Ala Val Ala Pro Val Leu Ser Ile
  1              5              10              15
Asn Ser Asp Phe Ser Asp Leu Arg Glu Ile Lys Lys Gln Leu Leu Leu
      20              25              30
Ile Ala Gly Leu Thr Arg Glu Arg Gly Leu Leu His Ser Ser Lys Trp
      35              40              45
Ser Ala Glu Leu Ala Phe Ser Leu Pro Ala Leu Pro Leu Ala Glu Leu
      50              55              60
Gln Pro Pro Pro Pro Ile Thr Glu Glu Asp Ala Gln Asp Met Asp Ala
      65              70              75              80
Tyr Thr Leu Ala Lys Ala Tyr Phe Asp Val Lys Glu Tyr Asp Arg Ala
      85              90              95
Ala His Phe Leu His Gly Cys Asn Ser Lys Lys Ala Tyr Phe Leu Tyr
      100              105              110
Met Tyr Ser Arg Tyr Leu Ser Gly Glu Lys Lys Lys Asp Asp Glu Thr
      115              120              125
Val Asp Ser Leu Gly Pro Leu Glu Lys Gly Gln Val Lys Asn Glu Ala
      130              135              140
Leu Arg Glu Leu Arg Val Glu Leu Ser Lys Lys His Gln Ala Arg Glu
      145              150              155              160
Leu Asp Gly Phe Gly Leu Tyr Leu Tyr Gly Val Val Leu Arg Lys Leu
      165              170              175
Asp Leu Val Lys Glu Ala Ile Asp Val Phe Val Glu Ala Thr His Val
      180              185              190
Leu Pro Leu His Trp Gly Ala Trp Leu Glu Leu Cys Asn Leu Ile Thr
      195              200              205
Asp Lys Glu Met Leu Lys Phe Leu Ser Leu Pro Asp Thr Trp Met Lys
      210              215              220

```

Glu Phe Phe Leu Ala His Ile Tyr Thr Glu Leu Gln Leu Ile Glu Glu
 225 230 235 240
 Ala Leu Gln Lys Tyr Gln Asn Leu Ile Asp Val Gly Phe Ser Lys Ser
 245 250 255
 Ser Tyr Ile Val Ser Gln Ile Ala Val Ala Tyr His Asn Ile Arg Asp
 260 265 270
 Ile Asp Lys Ala Leu Ser Ile Phe Asn Glu Leu Arg Lys Gln Asp Pro
 275 280 285
 Tyr Arg Ile Glu Asn Met Asp Thr Phe Ser Asn Leu Leu Tyr Val Arg
 290 295 300
 Ser Met Lys Ser Glu Leu Ser Tyr Leu Ala His Asn Leu Cys Glu Ile
 305 310 315 320
 Asp Lys Tyr Arg Val Glu Thr Cys Cys Val Ile Gly Asn Tyr Tyr Ser
 325 330 335
 Leu Arg Ser Gln His Glu Lys Ala Ala Leu Tyr Phe Gln Arg Ala Leu
 340 345 350
 Lys Leu Asn Pro Arg Tyr Leu Gly Ala Trp Thr Leu Met Gly His Glu
 355 360 365
 Tyr Met Glu Met Lys Asn Thr Ser Ala Ala Ile Gln Ala Tyr Arg His
 370 375 380
 Ala Ile Glu Val Asn Lys Arg Asp Tyr Arg Ala Trp Tyr Gly Leu Gly
 385 390 395 400
 Gln Thr Tyr Glu Ile Leu Lys Met Pro Phe Tyr Cys Leu Tyr Tyr Tyr
 405 410 415
 Arg Arg Ala His Gln Leu Arg Pro Asn Asp Ser Arg Met Leu Val Ala
 420 425 430
 Leu Gly Glu Cys Tyr Glu Lys Leu Asn Gln Leu Val Glu Ala Lys Lys
 435 440 445
 Cys Tyr Trp Arg Ala Tyr Ala Val Gly Asp Val Glu Lys Met Ala Leu
 450 455 460
 Val Lys Leu Ala Lys Leu His Glu Gln Leu Thr Glu Ser Glu Gln Ala
 465 470 475 480
 Ala Gln Cys Tyr Ile Lys Tyr Ile Gln Asp Ile Tyr Ser Cys Gly Glu
 485 490 495
 Ile Val Glu His Leu Glu Glu Ser Thr Ala Phe Arg Tyr Leu Ala Gln
 500 505 510
 Tyr Tyr Phe Lys Cys Lys Leu Trp Asp Glu Ala Ser Thr Cys Ala Gln
 515 520 525
 Lys Cys Cys Ala Phe Asn Asp Thr Arg Glu Glu Gly Lys Ala Leu Leu
 530 535 540
 Arg Gln Ile Leu Gln Leu Arg Asn Gln Gly Glu Thr Pro Thr Thr Glu
 545 550 555 560
 Val Pro Ala Pro Phe Phe Leu Pro Ala Ser Leu Ser Ala Asn Asn Thr
 565 570 575
 Pro Thr Arg Arg Val Ser Pro Leu Asn Leu Ser Ser Val Thr Pro
 580 585 590

<210> 672

<211> 914

<212> PRT

<213> Homo sapiens

<400> 672

Val Tyr Gln Val Leu Leu Val Gly Ser Thr Leu Leu Lys Glu Val Pro

1	5	10	15
Ser Gly Leu Gln Leu Glu Gln Leu Pro Ser Gln Ser Leu Leu Thr His			
20	25	30	
Ile Pro Thr Ala Gly Leu Pro Thr Ser Leu Gly Gly Gly Leu Pro Tyr			
35	40	45	
Cys His Gln Ala Trp Leu Asp Phe Arg Arg Arg Leu Glu Ala Leu Leu			
50	55	60	
Gln Asn Cys Gln Ala Ala Cys Ala Leu Leu Gln Gly Ala Ile Glu Ser			
65	70	75	80
Val Lys Ala Val Pro Gln Pro Met Glu Pro Gly Glu Val Gly Gln Leu			
85	90	95	
Leu Gln Gln Thr Glu Val Leu Met Gln Gln Val Leu Asp Ser Pro Trp			
100	105	110	
Leu Ala Trp Leu Gln Cys Gln Gly Gly Arg Glu Leu Thr Trp Leu Lys			
115	120	125	
Gln Glu Val Pro Glu Val Thr Leu Ser Pro Asp Tyr Arg Thr Ala Met			
130	135	140	
Asp Lys Ala Asp Glu Leu Tyr Asp Arg Val Asp Gly Leu Leu His Gln			
145	150	155	160
Leu Thr Leu Gln Ser Asn Gln Arg Ile Gln Ala Leu Glu Leu Val Gln			
165	170	175	
Thr Leu Glu Ala Arg Glu Ser Gly Leu His Gln Ile Glu Val Trp Leu			
180	185	190	
Gln Gln Val Gly Trp Pro Ala Leu Glu Glu Ala Gly Glu Pro Ser Leu			
195	200	205	
Asp Met Leu Leu Gln Ala Gln Gly Ser Phe Gln Glu Leu Tyr Gln Val			
210	215	220	
Ala Gln Glu Gln Val Arg Gln Gly Glu Lys Phe Leu Gln Pro Leu Thr			
225	230	235	240
Gly Trp Glu Ala Ala Glu Leu Asp Pro Pro Gly Ala Arg Phe Leu Ala			
245	250	255	
Leu Arg Ala Gln Leu Thr Glu Phe Ser Arg Ala Leu Ala Gln Arg Cys			
260	265	270	
Gln Arg Leu Ala Asp Ala Glu Arg Leu Phe Gln Leu Phe Arg Glu Ala			
275	280	285	
Leu Thr Trp Ala Glu Glu Gly Gln Arg Val Leu Ala Glu Leu Glu Gln			
290	295	300	
Glu Arg Pro Gly Val Val Leu Gln Gln Leu Gln Leu His Trp Thr Arg			
305	310	315	320
His Pro Asp Leu Pro Pro Ala His Phe Arg Lys Met Trp Ala Leu Ala			
325	330	335	
Thr Gly Leu Gly Ser Glu Ala Ile Arg Gln Glu Cys Arg Trp Ala Trp			
340	345	350	
Ala Arg Cys Gln Asp Thr Trp Leu Ala Leu Asp Gln Lys Leu Glu Ala			
355	360	365	
Ser Leu Lys Leu Pro Pro Val Gly Ser Thr Ala Ser Leu Cys Val Ser			
370	375	380	
Gln Val Pro Ala Ala Pro Ala His Pro Pro Leu Arg Lys Ala Tyr Ser			
385	390	395	400
Phe Asp Arg Asn Leu Gly Gln Ser Leu Ser Glu Pro Ala Cys His Cys			
405	410	415	
His His Ala Ala Thr Ile Ala Ala Cys Arg Arg Pro Glu Ala Gly Gly			
420	425	430	
Gly Ala Leu Pro Gln Ala Ser Pro Thr Val Pro Pro Pro Gly Ser Ser			

435					440					445					
Asp	Pro	Arg	Ser	Leu	Asn	Arg	Leu	Gln	Leu	Val	Leu	Ala	Glu	Met	Val
450					455					460					
Ala	Thr	Glu	Arg	Glu	Tyr	Val	Arg	Ala	Leu	Glu	Tyr	Thr	Met	Glu	Asn
465					470					475					
Tyr	Phe	Pro	Glu	Leu	Asp	Arg	Pro	Asp	Val	Pro	Gln	Gly	Leu	Arg	Gly
485					490					495					
Gln	Arg	Ala	His	Leu	Phe	Gly	Asn	Leu	Glu	Lys	Leu	Arg	Asp	Phe	His
500					505					510					
Cys	His	Phe	Phe	Leu	Arg	Glu	Leu	Glu	Ala	Cys	Thr	Arg	His	Pro	Pro
515					520					525					
Arg	Val	Ala	Tyr	Ala	Phe	Leu	Arg	His	Arg	Val	Gln	Phe	Gly	Met	Tyr
530					535					540					
Ala	Leu	Tyr	Ser	Lys	Asn	Lys	Pro	Arg	Ser	Asp	Ala	Leu	Met	Ser	Ser
545					550					555					
Tyr	Gly	His	Thr	Phe	Lys	Asp	Lys	Gln	Gln	Ala	Leu	Gly	Asp	His	
565					570					575					
Leu	Asp	Leu	Ala	Ser	Tyr	Leu	Leu	Lys	Pro	Ile	Gln	Arg	Met	Gly	Lys
580					585					590					
Tyr	Ala	Leu	Leu	Leu	Gln	Glu	Leu	Ala	Arg	Ala	Cys	Gly	Gly	Pro	Thr
595					600					605					
Gln	Glu	Leu	Ser	Ala	Leu	Arg	Glu	Ala	Gln	Ser	Leu	Val	His	Phe	Gln
610					615					620					
Leu	Arg	His	Gly	Asn	Asp	Leu	Leu	Ala	Met	Asp	Ala	Ile	Gln	Gly	Cys
625					630					635					
Asp	Val	Asn	Leu	Lys	Glu	Gln	Gly	Gln	Leu	Val	Arg	Gln	Asp	Glu	Phe
645					650					655					
Val	Val	Arg	Thr	Gly	Arg	His	Lys	Ser	Val	Arg	Arg	Ile	Phe	Leu	Phe
660					665					670					
Glu	Glu	Leu	Leu	Leu	Phe	Ser	Lys	Pro	Arg	His	Gly	Pro	Thr	Gly	Val
675					680					685					
Asp	Thr	Phe	Ala	Tyr	Lys	Arg	Ser	Phe	Lys	Met	Ala	Asp	Leu	Gly	Leu
690					695					700					
Thr	Glu	Cys	Cys	Gly	Asn	Ser	Asn	Leu	Arg	Phe	Glu	Ile	Trp	Phe	Arg
705					710					715					
Arg	Arg	Lys	Ala	Arg	Asp	Thr	Phe	Val	Leu	Gln	Ala	Ser	Ser	Leu	Ala
725					730					735					
Ile	Lys	Gln	Ala	Trp	Thr	Ala	Asp	Ile	Ser	His	Leu	Leu	Trp	Arg	Gln
740					745					750					
Ala	Val	His	Asn	Lys	Glu	Val	Arg	Met	Ala	Glu	Met	Val	Ser	Met	Gly
755					760					765					
Val	Gly	Asn	Lys	Ala	Phe	Arg	Asp	Ile	Ala	Pro	Ser	Glu	Glu	Ala	Ile
770					775					780					
Asn	Asp	Arg	Thr	Val	Asn	Tyr	Val	Leu	Lys	Cys	Arg	Glu	Val	Arg	Ser
785					790					795					
Arg	Ala	Ser	Ile	Ala	Val	Ala	Pro	Phe	Asp	His	Asp	Ser	Leu	Tyr	Leu
805					810					815					
Gly	Ala	Ser	Asn	Ser	Leu	Pro	Gly	Asp	Pro	Ala	Ser	Cys	Ser	Val	Leu
820					825					830					
Gly	Ser	Leu	Asn	Leu	His	Leu	Tyr	Arg	Asp	Pro	Ala	Leu	Leu	Gly	Leu
835					840					845					
Arg	Cys	Pro	Leu	Tyr	Pro	Ser	Phe	Leu	Glu	Glu	Ala	Ala	Leu	Glu	Ala
850					855					860					
Glu	Ala	Glu	Leu	Gly	Gly	Gln	Pro	Ser	Leu	Thr	Ala	Glu	Asp	Ser	Glu

865		870		875		880
Ile Ser Ser Gln Cys	Pro Ser Ala Ser Gly	Ser Ser Gly Ser	Asp Ser			
	885	890	895			
Ser Cys Val Ser Gly	Gln Ala Leu Gly	Arg Gly Leu Glu	Asp Leu Pro			
	900	905	910			
Cys Val						

<210> 673
 <211> 277
 <212> PRT
 <213> Homo sapiens

<400> 673

Leu Asn Tyr Leu Leu Glu Ser Arg Leu Glu Ala Ala Ala His Cys Ala																			
1		5		10		15													
Leu Lys Gln Gly Ile Ala Thr Ala Ser Leu Leu Pro Ala Gln Leu Gln																			
	20		25		30														
Pro Ala Val Leu Thr Val Val Thr Cys His Val Val Val Ser Val His																			
	35		40		45														
Gly His His Thr Asp Gly Cys Leu Ala Ala Leu Cys Arg Glu Asp Arg																			
	50		55		60														
Thr Gly Thr Gly Gly Ala Phe Trp Cys Lys Asn Arg Val Ile Val Ser																			
65		70		75		80													
His Ala Val Asp Val Val Leu His Val His Gly Glu Gly Asn Pro Val																			
	85		90		95														
Gln Ala Leu Ile Ala His Gly Ala Pro Glu Ala Ala Trp Val Val Gly																			
	100		105		110														
Leu Ala Gln Gly Leu Gln Asp His Phe His Asp Glu Met Ser Thr His																			
	115		120		125														
Ala Ala Phe Val Gly Arg Leu Leu Glu Pro Gly Val Gln Glu Val Leu																			
	130		135		140														
Leu Ala Val His Phe Leu Thr His Val Val Glu Arg Leu Pro Thr Glu																			
145		150		155		160													
Ser Ser Pro Thr Arg Val Ala Gly Glu Ala Val Ser Val Ile Lys Thr																			
	165		170		175														
Pro His Cys Leu Ala Arg Leu Leu Gly Ser Val Asp Ala Lys Pro Thr																			
	180		185		190														
Leu Asp Ala Asn Ala Glu Val Val Pro Arg Arg Ala Arg Leu Glu Arg																			
	195		200		205														
Pro Leu Gln Leu Pro Gly Glu Arg Leu Gln Pro Pro Leu Gly Arg Ala																			
	210		215		220														
Trp Ala Ala Leu Pro Ala Arg Gly Gln Arg Glu Cys Arg Gln Arg Glu																			
225		230		235		240													
Gly Gly Arg Pro Arg Arg Leu Arg Gly Ala Ser Gly Arg Gly Ala Gly																			
	245		250		255														
Ala Gly Arg Glu Glu Val Ser Val Gly Phe Ser Ala Gln Trp Glu Phe																			
	260		265		270														
Gly Ser Gly Arg His																			
	275																		

<210> 674
 <211> 1120
 <212> PRT

<213> Homo sapiens

<400> 674

```

Met Trp Arg Val Lys Lys Leu Ser Leu Ser Leu Ser Pro Ser Pro Gln
 1          5          10          15
Thr Gly Lys Pro Ser Met Arg Thr Pro Leu Arg Glu Leu Thr Leu Gln
          20          25          30
Pro Gly Ala Leu Thr Thr Ser Gly Lys Arg Ser Pro Ala Cys Ser Ser
          35          40          45
Leu Thr Pro Ser Leu Cys Lys Leu Gly Leu Gln Glu Gly Ser Asn Asn
          50          55          60
Ser Ser Pro Val Asp Phe Val Asn Asn Lys Arg Thr Asp Leu Ser Ser
65          70          75          80
Glu His Phe Ser His Ser Ser Lys Trp Leu Glu Thr Cys Gln His Glu
          85          90          95
Ser Asp Glu Gln Pro Leu Asp Pro Ile Pro Gln Ile Ser Ser Thr Pro
          100          105          110
Lys Thr Ser Glu Glu Ala Val Asp Pro Leu Gly Asn Tyr Met Val Lys
          115          120          125
Thr Ile Val Leu Val Pro Ser Pro Leu Gly Gln Gln Gln Asp Met Ile
          130          135          140
Phe Glu Ala Arg Leu Asp Thr Met Ala Glu Thr Asn Ser Ile Ser Leu
145          150          155          160
Asn Gly Pro Leu Arg Thr Asp Asp Leu Val Arg Glu Glu Val Ala Pro
          165          170          175
Cys Met Gly Asp Arg Phe Ser Glu Val Ala Ala Val Ser Glu Lys Pro
          180          185          190
Ile Phe Gln Glu Ser Pro Ser His Leu Leu Glu Glu Ser Pro Pro Asn
          195          200          205
Pro Cys Ser Glu Gln Leu His Cys Ser Lys Glu Ser Leu Ser Ser Arg
          210          215          220
Thr Glu Ala Val Arg Glu Asp Leu Val Pro Ser Glu Ser Asn Ala Phe
225          230          235          240
Leu Pro Ser Ser Val Leu Trp Leu Ser Pro Ser Thr Ala Leu Ala Ala
          245          250          255
Asp Phe Arg Val Asn His Val Asp Pro Glu Glu Glu Ile Val Glu His
          260          265          270
Gly Ala Met Glu Glu Arg Glu Met Arg Phe Pro Thr His Pro Lys Glu
          275          280          285
Ser Glu Thr Glu Asp Gln Ala Leu Val Ser Ser Val Glu Asp Ile Leu
          290          295          300
Ser Thr Cys Leu Thr Pro Asn Leu Val Glu Met Glu Ser Gln Glu Ala
305          310          315          320
Pro Gly Pro Ala Val Glu Asp Val Gly Arg Ile Leu Gly Ser Asp Thr
          325          330          335
Glu Ser Trp Met Ser Pro Leu Ala Trp Leu Glu Lys Gly Val Asn Thr
          340          345          350
Ser Val Met Leu Glu Asn Leu Arg Gln Ser Leu Ser Leu Pro Ser Met
          355          360          365
Leu Arg Asp Ala Ala Ile Gly Thr Thr Pro Phe Ser Thr Cys Ser Val
          370          375          380
Gly Thr Trp Phe Thr Pro Ser Ala Pro Gln Glu Lys Ser Thr Asn Thr
385          390          395          400
Ser Gln Thr Gly Leu Val Gly Thr Lys His Ser Thr Ser Glu Thr Glu

```

-232-

835	840	845
Thr Ala Lys Leu Ala Ser	Thr Ile Ala Asp Asn Gln Glu Gln Asp Leu	
850	855	860
Glu Lys Thr Arg Gln Tyr Ser	Gln Lys Leu Gly Leu Leu Thr Glu Gln	
865	870	875
Leu Gln Ser Leu Thr Leu Phe	Leu Gln Thr Lys Leu Lys Glu Lys Thr	
885	890	895
Glu Gln Glu Thr Leu Leu Leu Ser	Thr Ala Cys Pro Pro Thr Gln Glu	
900	905	910
His Pro Leu Pro Asn Asp Arg	Thr Phe Leu Gly Ser Ile Leu Thr Ala	
915	920	925
Val Ala Asp Glu Glu Pro Glu Ser	Thr Pro Val Pro Leu Leu Gly Ser	
930	935	940
Asp Lys Ser Ala Phe Thr Arg	Val Ala Ser Met Val Ser Leu Gln Pro	
945	950	955
Ala Glu Thr Pro Gly Met Glu Glu Ser	Leu Ala Glu Met Ser Ile Met	
965	970	975
Thr Thr Glu Leu Gln Ser Leu Cys Ser	Leu Leu Gln Glu Ser Lys Glu	
980	985	990
Glu Ala Ile Arg Thr Leu Gln Arg Lys	Ile Cys Glu Leu Gln Ala Arg	
995	1000	1005
Leu Gln Ala Gln Glu Glu Gln His Gln	Glu Val Gln Lys Ala Lys Glu	
1010	1015	1020
Ala Asp Ile Glu Lys Leu Asn Gln Ala	Leu Cys Leu Arg Tyr Lys Asn	
1025	1030	1035
Glu Lys Glu Leu Gln Glu Val Ile Gln	Gln Asn Glu Lys Ile Leu Glu	
1045	1050	1055
Gln Ile Asp Lys Ser Gly Glu Leu Ile Ser	Leu Arg Glu Glu Val Thr	
1060	1065	1070
His Leu Thr Arg Ser Leu Arg Arg Ala	Glu Thr Glu Thr Lys Val Leu	
1075	1080	1085
Gln Glu Ala Trp Gln Ala Ser Trp Thr	Pro Thr Ala Ser Leu Trp Pro	
1090	1095	1100
Pro Ile Gly Ser Arg Arg Lys Cys Gly Ser	Leu Arg Arg Trp Thr Asn	
1105	1110	1115
		1120

<210> 675

<211> 540

<212> PRT

<213> Homo sapiens

<400> 675

Met Gly Thr Thr Ala Arg Ala Ala Leu Val Leu Thr Tyr Leu Ala Val	
1 5 10 15	
Ala Ser Ala Ala Ser Glu Gly Gly Phe Thr Ala Thr Gly Gln Arg Gln	
20 25 30	
Leu Arg Pro Glu His Phe Gln Glu Val Gly Tyr Ala Ala Pro Pro Ser	
35 40 45	
Pro Pro Leu Ser Arg Ser Leu Pro Met Asp His Pro Asp Ser Ser Gln	
50 55 60	
His Gly Pro Pro Phe Glu Gly Gln Ser Gln Val Gln Pro Pro Pro Ser	
65 70 75 80	
Gln Glu Ala Thr Pro Leu Gln Gln Glu Lys Leu Leu Pro Ala Gln Leu	
85 90 95	

Pro	Ala	Glu	Lys	Glu	Val	Gly	Pro	Pro	Leu	Pro	Gln	Glu	Ala	Val	Pro	100	105	110
Leu	Gln	Lys	Glu	Leu	Pro	Ser	Leu	Gln	His	Pro	Asn	Glu	Gln	Lys	Glu	115	120	125
Gly	Thr	Pro	Ala	Pro	Phe	Gly	Asp	Gln	Ser	His	Pro	Glu	Pro	Glu	Ser	130	135	140
Trp	Asn	Ala	Ala	Gln	His	Cys	Gln	Gln	Asp	Arg	Ser	Gln	Gly	Gly	Trp	145	150	155
Gly	His	Arg	Leu	Asp	Gly	Phe	Pro	Pro	Gly	Arg	Pro	Ser	Pro	Asp	Asn	165	170	175
Leu	Asn	Gln	Ile	Cys	Leu	Pro	Asn	Arg	Gln	His	Val	Val	Tyr	Gly	Pro	180	185	190
Trp	Asn	Leu	Pro	Gln	Ser	Ser	Tyr	Ser	His	Leu	Thr	Arg	Gln	Gly	Glu	195	200	205
Thr	Leu	Asn	Phe	Leu	Glu	Ile	Gly	Tyr	Ser	Arg	Cys	Cys	His	Cys	Arg	210	215	220
Ser	His	Thr	Asn	Arg	Leu	Glu	Cys	Ala	Lys	Leu	Val	Trp	Glu	Glu	Ala	225	230	235
Met	Ser	Arg	Phe	Cys	Glu	Ala	Glu	Phe	Ser	Val	Lys	Thr	Arg	Pro	His	245	250	255
Trp	Cys	Cys	Thr	Arg	Gln	Gly	Glu	Ala	Arg	Phe	Ser	Cys	Phe	Gln	Glu	260	265	270
Glu	Ala	Pro	Gln	Pro	His	Tyr	Gln	Leu	Arg	Ala	Cys	Pro	Ser	His	Gln	275	280	285
Pro	Asp	Ile	Ser	Ser	Gly	Leu	Glu	Leu	Pro	Phe	Pro	Pro	Gly	Val	Pro	290	295	300
Thr	Leu	Asp	Asn	Ile	Lys	Asn	Ile	Cys	His	Leu	Arg	Arg	Phe	Arg	Ser	305	310	315
Val	Pro	Arg	Asn	Leu	Pro	Ala	Thr	Asp	Pro	Leu	Gln	Arg	Glu	Leu	Leu	325	330	335
Ala	Leu	Ile	Gln	Leu	Glu	Arg	Glu	Phe	Gln	Arg	Cys	Cys	Arg	Gln	Gly	340	345	350
Asn	Asn	His	Thr	Cys	Thr	Trp	Lys	Ala	Trp	Glu	Asp	Thr	Leu	Asp	Lys	355	360	365
Tyr	Cys	Asp	Arg	Glu	Tyr	Ala	Val	Lys	Thr	His	His	His	Leu	Cys	Cys	370	375	380
Arg	His	Pro	Pro	Ser	Pro	Thr	Arg	Asp	Glu	Cys	Phe	Ala	Arg	Arg	Ala	385	390	395
Pro	Tyr	Pro	Asn	Tyr	Asp	Arg	Asp	Ile	Leu	Thr	Ile	Asp	Ile	Ser	Arg	405	410	415
Val	Thr	Pro	Asn	Leu	Met	Gly	His	Leu	Cys	Gly	Asn	Gln	Arg	Val	Leu	420	425	430
Thr	Lys	His	Lys	His	Ile	Pro	Gly	Leu	Ile	His	Asn	Met	Thr	Ala	Arg	435	440	445
Cys	Cys	Asp	Leu	Pro	Phe	Pro	Glu	Gln	Ala	Cys	Cys	Ala	Glu	Glu	Glu	450	455	460
Lys	Leu	Thr	Phe	Ile	Asn	Asp	Leu	Cys	Gly	Pro	Arg	Arg	Asn	Ile	Trp	465	470	475
Arg	Asp	Pro	Ala	Leu	Cys	Cys	Tyr	Leu	Ser	Pro	Gly	Asp	Glu	Gln	Val	485	490	495
Asn	Cys	Phe	Asn	Ile	Asn	Tyr	Leu	Arg	Asn	Val	Ala	Leu	Val	Ser	Gly	500	505	510
Asp	Thr	Glu	Asn	Ala	Lys	Gly	Gln	Gly	Glu	Gln	Gly	Ser	Thr	Gly	Gly	515	520	525

Thr Asn Ile Ser Ser Thr Ser Glu Pro Lys Glu Glu
 530 535 540

<210> 676

<211> 462

<212> PRT

<213> Homo sapiens

<400> 676

Met	Gly	Lys	Glu	Lys	Thr	His	Ile	Asn	Ile	Val	Val	Ile	Gly	His	Val
1				5					10					15	
Asp	Ser	Gly	Lys	Ser	Thr	Thr	Thr	Gly	His	Leu	Ile	Tyr	Lys	Cys	Gly
			20					25					30		
Gly	Ile	Asp	Lys	Arg	Thr	Ile	Glu	Lys	Phe	Glu	Lys	Glu	Ala	Ala	Glu
		35					40					45			
Met	Gly	Lys	Gly	Ser	Phe	Lys	Tyr	Ala	Trp	Val	Leu	Asp	Lys	Leu	Lys
50						55					60				
Ala	Glu	Arg	Glu	Arg	Gly	Ile	Thr	Ile	Asp	Ile	Ser	Leu	Trp	Lys	Phe
65					70					75					80
Glu	Thr	Ser	Lys	Tyr	Tyr	Val	Thr	Ile	Ile	Asp	Ala	Pro	Gly	His	Arg
				85					90					95	
Asp	Phe	Ile	Lys	Asn	Met	Ile	Thr	Gly	Thr	Ser	Gln	Ala	Asp	Cys	Ala
			100					105					110		
Val	Leu	Ile	Val	Ala	Ala	Gly	Val	Gly	Glu	Phe	Glu	Ala	Gly	Ile	Ser
		115					120					125			
Lys	Asn	Gly	Gln	Thr	Arg	Glu	His	Ala	Leu	Leu	Ala	Tyr	Thr	Leu	Gly
		130				135					140				
Val	Lys	Gln	Leu	Ile	Val	Gly	Val	Asn	Lys	Met	Asp	Ser	Thr	Glu	Pro
145					150					155					160
Pro	Tyr	Ser	Gln	Lys	Arg	Tyr	Glu	Glu	Ile	Val	Lys	Glu	Val	Ser	Thr
				165					170					175	
Tyr	Ile	Lys	Lys	Ile	Gly	Tyr	Asn	Pro	Asp	Thr	Val	Ala	Phe	Val	Pro
		180						185					190		
Ile	Ser	Gly	Trp	Asn	Gly	Asp	Asn	Met	Leu	Glu	Pro	Ser	Ala	Asn	Met
		195				200						205			
Pro	Trp	Phe	Lys	Gly	Trp	Lys	Val	Thr	Arg	Lys	Asp	Gly	Asn	Ala	Ser
	210					215					220				
Gly	Thr	Thr	Leu	Leu	Glu	Ala	Val	Asp	Cys	Ile	Leu	Pro	Pro	Thr	Arg
225					230					235					240
Pro	Thr	Asp	Lys	Pro	Leu	Arg	Leu	Pro	Leu	Gln	Asp	Val	Tyr	Lys	Ile
				245						250				255	
Gly	Gly	Ile	Gly	Thr	Val	Pro	Val	Gly	Arg	Val	Glu	Thr	Gly	Val	Leu
		260						265					270		
Lys	Pro	Gly	Met	Val	Val	Thr	Phe	Ala	Pro	Val	Asn	Val	Thr	Thr	Glu
		275					280					285			
Val	Lys	Ser	Val	Glu	Met	His	His	Glu	Ala	Leu	Ser	Glu	Ala	Leu	Pro
		290				295					300				
Gly	Asp	Asn	Val	Gly	Phe	Asn	Val	Lys	Asn	Val	Ser	Val	Lys	Asp	Val
305				310						315					320
Arg	Arg	Gly	Asn	Val	Ala	Gly	Asp	Ser	Lys	Asn	Asp	Pro	Pro	Met	Glu
				325						330				335	
Ala	Ala	Gly	Phe	Thr	Ala	Gln	Val	Ile	Ile	Leu	Asn	His	Pro	Gly	Gln
		340						345					350		
Ile	Ser	Ala	Gly	Tyr	Ala	Pro	Val	Leu	Asp	Cys	His	Thr	Ala	His	Ile

355	360	365
Ala Cys Lys Phe	Ala Glu Leu Lys Glu Lys Ile	Asp Arg Arg Ser Gly
370	375	380
Lys Lys Leu Glu	Asp Gly Pro Lys Phe Leu Lys	Ser Gly Asp Ala Ala
385	390	395
Ile Val Asp Met	Val Pro Gly Lys Pro Met Cys Val	Glu Ser Phe Ser
405	410	415
Asp Tyr Pro Pro	Leu Gly Arg Phe Ala Val Arg	Asp Met Arg Gln Thr
420	425	430
Val Ala Val Gly	Val Ile Lys Ala Val Asp Lys Lys	Ala Ala Gly Ala
435	440	445
Gly Lys Val Thr	Lys Ser Ala Gln Lys Ala Gln	Lys Lys Ala Lys
450	455	460

<210> 677

<211> 2328

<212> PRT

<213> Homo sapiens

<400> 677

Lys Ser Lys Arg	Gln Ala Gln Gln Met	Val Gln Pro Gln Ser	Pro Val
1	5	10	15
Ala Val Ser Gln	Ser Lys Pro Gly Cys Tyr	Asp Asn Gly Lys	His Tyr
20	25	30	
Gln Ile Asn Gln	Gln Trp Glu Arg Thr Tyr	Leu Gly Asn Val	Leu Val
35	40	45	
Cys Thr Cys Tyr	Gly Gly Ser Arg Gly Phe	Asn Cys Glu Ser	Lys Pro
50	55	60	
Glu Ala Glu Glu	Thr Cys Phe Asp Lys Tyr	Thr Gly Asn Thr	Tyr Arg
65	70	75	80
Val Gly Asp Thr	Tyr Glu Arg Pro Lys Asp	Ser Met Ile Trp	Asp Cys
85	90	95	
Thr Cys Ile Gly	Ala Gly Arg Gly Arg Ile	Ser Cys Thr Ile	Ala Asn
100	105	110	
Arg Cys His Glu	Gly Gly Gln Ser Tyr Lys	Ile Gly Asp Thr	Trp Arg
115	120	125	
Arg Pro His Glu	Thr Gly Gly Tyr Met Leu	Glu Cys Val Cys	Leu Gly
130	135	140	
Asn Gly Lys Gly	Glu Trp Thr Cys Lys Pro	Ile Ala Glu Lys	Cys Phe
145	150	155	160
Asp His Ala Ala	Gly Thr Ser Tyr Val Val	Gly Glu Thr Trp	Glu Lys
165	170	175	
Pro Tyr Gln Gly	Trp Met Met Val Asp Cys	Thr Cys Leu Gly	Glu Gly
180	185	190	
Ser Gly Arg Ile	Thr Cys Thr Ser Arg Asn	Arg Cys Asn Asp	Gln Asp
195	200	205	
Thr Arg Thr Ser	Tyr Arg Ile Gly Asp Thr	Trp Ser Lys Lys	Asp Asn
210	215	220	
Arg Gly Asn Leu	Leu Gln Cys Ile Cys Thr	Gly Asn Gly Arg	Gly Glu
225	230	235	240
Trp Lys Cys Glu	Arg His Thr Ser Val Gln	Thr Thr Ser Ser	Gly Ser
245	250	255	
Gly Pro Phe Thr	Asp Val Arg Ala Ala Val	Tyr Gln Pro Gln	Pro His
260	265	270	

Pro	Gln	Pro	Pro	Pro	Tyr	Gly	His	Cys	Val	Thr	Asp	Ser	Gly	Val	Val
		275					280					285			
Tyr	Ser	Val	Gly	Met	Gln	Trp	Leu	Lys	Thr	Gln	Gly	Asn	Lys	Gln	Met
		290				295					300				
Leu	Cys	Thr	Cys	Leu	Gly	Asn	Gly	Val	Ser	Cys	Gln	Glu	Thr	Ala	Val
305					310					315					320
Thr	Gln	Thr	Tyr	Gly	Gly	Asn	Leu	Asn	Gly	Glu	Pro	Cys	Val	Leu	Pro
				325					330					335	
Phe	Thr	Tyr	Asn	Gly	Arg	Thr	Phe	Tyr	Ser	Cys	Thr	Thr	Glu	Gly	Arg
			340					345					350		
Gln	Asp	Gly	His	Leu	Trp	Cys	Ser	Thr	Thr	Ser	Asn	Tyr	Glu	Gln	Asp
		355					360					365			
Gln	Lys	Tyr	Ser	Phe	Cys	Thr	Asp	His	Thr	Val	Leu	Val	Gln	Thr	Gln
		370				375					380				
Gly	Gly	Asn	Ser	Asn	Gly	Ala	Leu	Cys	His	Phe	Pro	Phe	Leu	Tyr	Asn
385					390					395					400
Asn	His	Asn	Tyr	Thr	Asp	Cys	Thr	Ser	Glu	Gly	Arg	Arg	Asp	Asn	Met
				405					410					415	
Lys	Trp	Cys	Gly	Thr	Thr	Gln	Asn	Tyr	Asp	Ala	Asp	Gln	Lys	Phe	Gly
			420					425					430		
Phe	Cys	Pro	Met	Ala	Ala	His	Glu	Glu	Ile	Cys	Thr	Thr	Asn	Glu	Gly
		435					440					445			
Val	Met	Tyr	Arg	Ile	Gly	Asp	Gln	Trp	Asp	Lys	Gln	His	Asp	Met	Gly
		450				455					460				
His	Met	Met	Arg	Cys	Thr	Cys	Val	Gly	Asn	Gly	Arg	Gly	Glu	Trp	Thr
465					470					475					480
Cys	Ile	Ala	Tyr	Ser	Gln	Leu	Arg	Asp	Gln	Cys	Ile	Val	Asp	Asp	Ile
				485				490						495	
Thr	Tyr	Asn	Val	Asn	Asp	Thr	Phe	His	Lys	Arg	His	Glu	Glu	Gly	His
			500					505					510		
Met	Leu	Asn	Cys	Thr	Cys	Phe	Gly	Gln	Gly	Arg	Gly	Arg	Trp	Lys	Cys
		515					520					525			
Asp	Pro	Val	Asp	Gln	Cys	Gln	Asp	Ser	Glu	Thr	Gly	Thr	Phe	Tyr	Gln
		530				535					540				
Ile	Gly	Asp	Ser	Trp	Glu	Lys	Tyr	Val	His	Gly	Val	Arg	Tyr	Gln	Cys
545					550					555					560
Tyr	Cys	Tyr	Gly	Arg	Gly	Ile	Gly	Glu	Trp	His	Cys	Gln	Pro	Leu	Gln
				565				570						575	
Thr	Tyr	Pro	Ser	Ser	Ser	Gly	Pro	Val	Glu	Val	Phe	Ile	Thr	Glu	Thr
			580					585					590		
Pro	Ser	Gln	Pro	Asn	Ser	His	Pro	Ile	Gln	Trp	Asn	Ala	Pro	Gln	Pro
		595					600					605			
Ser	His	Ile	Ser	Lys	Tyr	Ile	Leu	Arg	Trp	Arg	Pro	Lys	Asn	Ser	Val
		610				615					620				
Gly	Arg	Trp	Lys	Glu	Ala	Thr	Ile	Pro	Gly	His	Leu	Asn	Ser	Tyr	Thr
625					630					635					640
Ile	Lys	Gly	Leu	Lys	Pro	Gly	Val	Val	Tyr	Glu	Gly	Gln	Leu	Ile	Ser
				645				650						655	
Ile	Gln	Gln	Tyr	Gly	His	Gln	Glu	Val	Thr	Arg	Phe	Asp	Phe	Thr	Thr
			660					665					670		
Thr	Ser	Thr	Ser	Thr	Pro	Val	Thr	Ser	Asn	Thr	Val	Thr	Gly	Glu	Thr
		675					680					685			
Thr	Pro	Phe	Ser	Pro	Leu	Val	Ala	Thr	Ser	Glu	Ser	Val	Thr	Glu	Ile
		690				695					700				

Thr	Ala	Ser	Ser	Phe	Val	Val	Ser	Trp	Val	Ser	Ala	Ser	Asp	Thr	Val	705	710	715	720
Ser	Gly	Phe	Arg	Val	Glu	Tyr	Glu	Leu	Ser	Glu	Glu	Gly	Asp	Glu	Pro	725	730	735	
Gln	Tyr	Leu	Asp	Leu	Pro	Ser	Thr	Ala	Thr	Ser	Val	Asn	Ile	Pro	Asp	740	745	750	
Leu	Leu	Pro	Gly	Arg	Lys	Tyr	Ile	Val	Asn	Val	Tyr	Gln	Ile	Ser	Glu	755	760	765	
Asp	Gly	Glu	Gln	Ser	Leu	Ile	Leu	Ser	Thr	Ser	Gln	Thr	Thr	Ala	Pro	770	775	780	
Asp	Ala	Pro	Pro	Asp	Pro	Thr	Val	Asp	Gln	Val	Asp	Asp	Thr	Ser	Ile	785	790	795	800
Val	Val	Arg	Trp	Ser	Arg	Pro	Gln	Ala	Pro	Ile	Thr	Gly	Tyr	Arg	Ile	805	810		815
Val	Tyr	Ser	Pro	Ser	Val	Glu	Gly	Ser	Ser	Thr	Glu	Leu	Asn	Leu	Pro	820	825	830	
Glu	Thr	Ala	Asn	Ser	Val	Thr	Leu	Ser	Asp	Leu	Gln	Pro	Gly	Val	Gln	835	840	845	
Tyr	Asn	Ile	Thr	Ile	Tyr	Ala	Val	Glu	Glu	Asn	Gln	Glu	Ser	Thr	Pro	850	855	860	
Val	Val	Ile	Gln	Gln	Glu	Thr	Thr	Gly	Thr	Pro	Arg	Ser	Asp	Thr	Val	865	870	875	880
Pro	Ser	Pro	Arg	Asp	Leu	Gln	Phe	Val	Glu	Val	Thr	Asp	Val	Lys	Val	885	890		895
Thr	Ile	Met	Trp	Thr	Pro	Pro	Glu	Ser	Ala	Val	Thr	Gly	Tyr	Arg	Val	900	905	910	
Asp	Val	Ile	Pro	Val	Asn	Leu	Pro	Gly	Glu	His	Gly	Gln	Arg	Leu	Pro	915	920	925	
Ile	Ser	Arg	Asn	Thr	Phe	Ala	Glu	Val	Thr	Gly	Leu	Ser	Pro	Gly	Val	930	935	940	
Thr	Tyr	Tyr	Phe	Lys	Val	Phe	Ala	Val	Ser	His	Gly	Arg	Glu	Ser	Lys	945	950	955	960
Pro	Leu	Thr	Ala	Gln	Gln	Thr	Thr	Lys	Leu	Asp	Ala	Pro	Thr	Asn	Leu	965	970		975
Gln	Phe	Val	Asn	Glu	Thr	Asp	Ser	Thr	Val	Leu	Val	Arg	Trp	Thr	Pro	980	985	990	
Pro	Arg	Ala	Gln	Ile	Thr	Gly	Tyr	Arg	Leu	Thr	Val	Gly	Leu	Thr	Arg	995	1000	1005	
Arg	Gly	Gln	Pro	Arg	Gln	Tyr	Asn	Val	Gly	Pro	Ser	Val	Ser	Lys	Tyr	1010	1015	1020	
Pro	Leu	Arg	Asn	Leu	Gln	Pro	Ala	Ser	Glu	Tyr	Thr	Val	Ser	Leu	Val	1025	1030	1035	1040
Ala	Ile	Lys	Gly	Asn	Gln	Glu	Ser	Pro	Lys	Ala	Thr	Gly	Val	Phe	Thr	1045	1050		1055
Thr	Leu	Gln	Pro	Gly	Ser	Ser	Ile	Pro	Pro	Tyr	Asn	Thr	Glu	Val	Thr	1060	1065	1070	
Glu	Thr	Thr	Ile	Val	Ile	Thr	Trp	Thr	Pro	Ala	Pro	Arg	Ile	Gly	Phe	1075	1080	1085	
Lys	Leu	Gly	Val	Arg	Pro	Ser	Gln	Gly	Gly	Glu	Ala	Pro	Arg	Glu	Val	1090	1095	1100	
Thr	Ser	Asp	Ser	Gly	Ser	Ile	Val	Val	Ser	Gly	Leu	Thr	Pro	Gly	Val	1105	1110	1115	1120
Glu	Tyr	Val	Tyr	Thr	Ile	Gln	Val	Leu	Arg	Asp	Gly	Gln	Glu	Arg	Asp	1125	1130		1135

Ala Pro Ile Val Asn Lys Val Val Thr Pro Leu Ser Pro Pro Thr Asn
 1140 1145 1150
 Leu His Leu Glu Ala Asn Pro Asp Thr Gly Val Leu Thr Val Ser Trp
 1155 1160 1165
 Glu Arg Ser Thr Thr Pro Asp Ile Thr Gly Tyr Arg Ile Thr Thr Thr
 1170 1175 1180
 Pro Thr Asn Gly Gln Gln Gly Asn Ser Leu Glu Glu Val Val His Ala
 1185 1190 1195 1200
 Asp Gln Ser Ser Cys Thr Phe Asp Asn Leu Ser Pro Gly Leu Glu Tyr
 1205 1210 1215
 Asn Val Ser Val Tyr Thr Val Lys Asp Asp Lys Glu Ser Val Pro Ile
 1220 1225 1230
 Ser Asp Thr Ile Ile Pro Ala Val Pro Pro Pro Thr Asp Leu Arg Phe
 1235 1240 1245
 Thr Asn Ile Gly Pro Asp Thr Met Arg Val Thr Trp Ala Pro Pro Pro
 1250 1255 1260
 Ser Ile Asp Leu Thr Asn Phe Leu Val Arg Tyr Ser Pro Val Lys Asn
 1265 1270 1275 1280
 Glu Glu Asp Val Ala Glu Leu Ser Ile Ser Pro Ser Asp Asn Ala Val
 1285 1290 1295
 Val Leu Thr Asn Leu Leu Pro Gly Thr Glu Tyr Val Val Ser Val Ser
 1300 1305 1310
 Ser Val Tyr Glu Gln His Glu Ser Thr Pro Leu Arg Gly Arg Gln Lys
 1315 1320 1325
 Thr Gly Leu Asp Ser Pro Thr Gly Ile Asp Phe Ser Asp Ile Thr Ala
 1330 1335 1340
 Asn Ser Phe Thr Val His Trp Ile Ala Pro Arg Ala Thr Ile Thr Gly
 1345 1350 1355 1360
 Tyr Arg Ile Arg His His Pro Glu His Phe Ser Gly Arg Pro Arg Glu
 1365 1370 1375
 Asp Arg Val Pro His Ser Arg Asn Ser Ile Thr Leu Thr Asn Leu Thr
 1380 1385 1390
 Pro Gly Thr Glu Tyr Val Val Ser Ile Val Ala Leu Asn Gly Arg Glu
 1395 1400 1405
 Glu Ser Pro Leu Leu Ile Gly Gln Gln Ser Thr Val Ser Asp Val Pro
 1410 1415 1420
 Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu Ile Ser
 1425 1430 1435 1440
 Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly
 1445 1450 1455
 Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe Thr Val Pro Gly Ser
 1460 1465 1470
 Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro Gly Val Asp Tyr Thr
 1475 1480 1485
 Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp Ser Pro Ala Ser Ser
 1490 1495 1500
 Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln
 1505 1510 1515 1520
 Met Gln Val Thr Asp Val Gln Asp Asn Ser Ile Ser Val Lys Trp Leu
 1525 1530 1535
 Pro Ser Ser Ser Pro Val Thr Gly Tyr Arg Val Thr Thr Thr Pro Lys
 1540 1545 1550
 Asn Gly Pro Gly Pro Thr Lys Thr Lys Thr Ala Gly Pro Asp Gln Thr
 1555 1560 1565

-240-

Tyr Asp Thr Gly Asn Gly Ile Gln Leu Pro Gly Thr Ser Gly Gln Gln
 2005 2010 2015
 Pro Ser Val Gly Gln Gln Met Ile Phe Glu Glu His Gly Phe Arg Arg
 2020 2025 2030
 Thr Thr Pro Pro Thr Thr Ala Thr Pro Ile Arg His Arg Pro Arg Pro
 2035 2040 2045
 Tyr Pro Pro Asn Val Gly Gln Glu Ala Leu Ser Gln Thr Thr Ile Ser
 2050 2055 2060
 Trp Ala Pro Phe Gln Asp Thr Ser Glu Tyr Ile Ile Ser Cys His Pro
 2065 2070 2075 2080
 Val Gly Thr Asp Glu Glu Pro Leu Gln Phe Arg Val Pro Gly Thr Ser
 2085 2090 2095
 Thr Ser Ala Thr Leu Thr Gly Leu Thr Arg Gly Ala Thr Tyr Asn Ile
 2100 2105 2110
 Ile Val Glu Ala Leu Lys Asp Gln Gln Arg His Lys Val Arg Glu Glu
 2115 2120 2125
 Val Val Thr Val Gly Asn Ser Val Asn Glu Gly Leu Asn Gln Pro Thr
 2130 2135 2140
 Asp Asp Ser Cys Phe Asp Pro Tyr Thr Val Ser His Tyr Ala Val Gly
 2145 2150 2155 2160
 Asp Glu Trp Glu Arg Met Ser Glu Ser Gly Phe Lys Leu Leu Cys Gln
 2165 2170 2175
 Cys Leu Gly Phe Gly Ser Gly His Phe Arg Cys Asp Ser Ser Arg Trp
 2180 2185 2190
 Cys His Asp Asn Gly Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg
 2195 2200 2205
 Gln Gly Glu Asn Gly Gln Met Met Ser Cys Thr Cys Leu Gly Asn Gly
 2210 2215 2220
 Lys Gly Glu Phe Lys Cys Asp Pro His Glu Ala Thr Cys Tyr Asp Asp
 2225 2230 2235 2240
 Gly Lys Thr Tyr His Val Gly Glu Gln Trp Gln Lys Glu Tyr Leu Gly
 2245 2250 2255
 Ala Ile Cys Ser Cys Thr Cys Phe Gly Gly Gln Arg Gly Trp Arg Cys
 2260 2265 2270
 Asp Asn Cys Arg Arg Pro Gly Gly Glu Pro Ser Pro Glu Gly Thr Thr
 2275 2280 2285
 Gly Gln Ser Tyr Asn Gln Tyr Ser Gln Arg Tyr His Gln Arg Thr Asn
 2290 2295 2300
 Thr Asn Val Asn Cys Pro Ile Glu Cys Phe Met Pro Leu Asp Val Gln
 2305 2310 2315 2320
 Ala Asp Arg Glu Asp Ser Arg Glu
 2325

<210> 678

<211> 188

<212> PRT

<213> Homo sapiens

<400> 678

His Gln Thr His Lys Glu Gly Gly Ser Thr His Ala Ser Ala Asp Ala
 1 5 10 15
 Trp Glu Ile Ile Glu Leu Glu Thr Glu Ile Glu Lys Phe Lys Ala Glu
 20 25 30
 Asn Ala Ser Leu Ala Lys Leu Arg Ile Glu Arg Glu Ser Ala Leu Glu

```
<210> 679
<211> 284
<212> PRT
<213> Homo sapiens
```

-242-

Lys Gly Gln Leu Arg Glu Leu Glu Arg Glu Tyr Ala Ala Asn Lys Phe
 225 230 235 240
 Ile Thr Lys Asp Lys Arg Arg Lys Ile Ser Ala Ala Thr Ser Leu Ser
 245 250 255
 Glu Arg Gln Ile Thr Ile Trp Phe Gln Asn Arg Arg Val Lys Glu Lys
 260 265 270
 Lys Val Leu Ala Lys Val Lys Asn Ser Ala Thr Pro
 275 280

<210> 680

<211> 676

<212> PRT

<213> Homo sapiens

<400> 680

Met Asp Lys Tyr Asp Asp Leu Gly Leu Glu Ala Ser Lys Phe Ile Glu
 1 5 10 15
 Asp Leu Asn Met Tyr Glu Ala Ser Lys Asp Gly Leu Phe Arg Val Asp
 20 25 30
 Lys Gly Ala Gly Asn Asn Pro Glu Phe Glu Glu Thr Arg Arg Val Phe
 35 40 45
 Ala Thr Lys Met Ala Lys Ile His Leu Gln Gln Gln Gln Gln Leu
 50 55 60
 Leu Gln Glu Glu Thr Leu Pro Arg Gly Ser Arg Gly Pro Val Asn Gly
 65 70 75 80
 Gly Gly Arg Leu Gly Pro Gln Ala Arg Trp Glu Val Val Gly Ser Lys
 85 90 95
 Leu Thr Val Asp Gly Ala Ala Lys Pro Pro Leu Ala Ala Ser Thr Gly
 100 105 110
 Ala Pro Gly Ala Val Thr Thr Leu Ala Ala Gly Gln Pro Pro Tyr Pro
 115 120 125
 Pro Gln Glu Gln Arg Ser Arg Pro Tyr Leu His Gly Thr Arg His Gly
 130 135 140
 Ser Gln Asp Cys Gly Ser Arg Glu Ser Leu Ala Thr Ser Glu Met Ser
 145 150 155 160
 Ala Phe His Gln Pro Gly Pro Cys Glu Asp Pro Ser Cys Leu Thr His
 165 170 175
 Gly Asp Tyr Tyr Asp Asn Leu Ser Leu Ala Ser Pro Lys Trp Gly Asp
 180 185 190
 Lys Pro Gly Val Ser Pro Ser Ile Gly Leu Ser Val Gly Ser Gly Trp
 195 200 205
 Pro Ser Ser Pro Gly Ser Asp Pro Pro Leu Pro Lys Pro Cys Gly Asp
 210 215 220
 His Pro Leu Asn His Arg Gln Leu Ser Leu Ser Ser Ser Arg Ser Ser
 225 230 235 240
 Glu Gly Ser Leu Gly Gly Gln Asn Ser Gly Ile Gly Gly Arg Ser Ser
 245 250 255
 Glu Lys Pro Thr Gly Leu Trp Ser Thr Ala Ser Ser Gln Arg Val Ser
 260 265 270
 Pro Gly Leu Pro Ser Pro Asn Leu Glu Asn Gly Ala Pro Ala Val Gly
 275 280 285
 Pro Val Gln Pro Arg Thr Pro Ser Val Ser Ala Pro Leu Ala Leu Ser
 290 295 300
 Cys Pro Arg Gln Gly Gly Leu Pro Arg Ser Asn Ser Gly Leu Gly Gly

305 310 315 320
 Glu Val Ser Gly Val Met Ser Lys Pro Asn Val Asp Pro Gln Pro Trp
 325 330 335
 Phe Gln Asp Gly Pro Lys Ser Tyr Leu Ser Ser Ser Ala Pro Ser Ser
 340 345 350
 Ser Pro Ala Gly Leu Asp Gly Ser Gln Gln Gly Ala Val Pro Gly Leu
 355 360 365
 Gly Pro Lys Pro Gly Cys Thr Asp Leu Gly Thr Gly Pro Lys Leu Ser
 370 375 380
 Pro Thr Ser Leu Val His Pro Val Met Ser Thr Leu Pro Glu Leu Ser
 385 390 395 400
 Cys Lys Glu Gly Pro Leu Gly Trp Ser Ser Asp Gly Ser Leu Gly Ser
 405 410 415
 Val Leu Leu Asp Ser Pro Ser Ser Pro Arg Val Arg Leu Pro Cys Gln
 420 425 430
 Pro Leu Val Pro Gly Pro Glu Leu Arg Pro Ser Ala Ala Glu Leu Lys
 435 440 445
 Leu Glu Ala Leu Thr Gln Arg Leu Glu Arg Glu Met Asp Ala His Pro
 450 455 460
 Lys Ala Asp Tyr Phe Gly Ala Cys Val Lys Cys Ser Lys Gly Val Phe
 465 470 475 480
 Gly Ala Gly Gln Ala Cys Gln Ala Met Gly Asn Leu Tyr His Asp Thr
 485 490 495
 Cys Phe Thr Cys Ala Ala Cys Ser Arg Lys Leu Arg Gly Lys Ala Phe
 500 505 510
 Tyr Phe Val Asn Gly Lys Val Phe Cys Glu Glu Asp Phe Leu Tyr Ser
 515 520 525
 Gly Phe Gln Gln Ser Ala Asp Arg Cys Phe Leu Cys Gly His Leu Ile
 530 535 540
 Met Asp Met Ile Leu Gln Ala Leu Gly Lys Ser Tyr His Pro Gly Cys
 545 550 555 560
 Phe Arg Cys Val Ile Cys Asn Glu Cys Leu Asp Gly Val Pro Phe Thr
 565 570 575
 Val Asp Ser Glu Asn Lys Ile Tyr Cys Val Arg Asp Tyr His Lys Val
 580 585 590
 Leu Ala Pro Lys Cys Ala Ala Cys Gly Leu Pro Ile Leu Pro Pro Glu
 595 600 605
 Gly Ser Asp Glu Thr Ile Arg Val Val Ser Met Asp Arg Asp Tyr His
 610 615 620
 Val Glu Cys Tyr His Cys Glu Asp Cys Gly Leu Glu Leu Asn Asp Glu
 625 630 635 640
 Asp Gly His Arg Cys Tyr Pro Leu Glu Asp His Leu Phe Cys His Ser
 645 650 655
 Cys His Val Lys Arg Leu Glu Lys Arg Pro Ser Ser Thr Ala Leu His
 660 665 670
 Gln His His Phe
 675

<210> 681
 <211> 296
 <212> PRT
 <213> Homo sapiens

<400> 681

```

Ser Thr Gly Ser Glu Phe Pro Leu Cys Thr Lys Ala Ser Pro Cys Ser
 1          5          10          15
Ala Ala Arg Ala Gly Gly Arg Ala Leu Gly Trp Arg Leu Gln Gln Gln
          20          25          30
Arg Glu Thr Arg Gly Asn Pro Gly Asn Pro Gly Leu Gly Val Ala Ala
          35          40          45
Thr Met Thr Gly Ser Asn Met Ser Asp Ala Leu Ala Asn Ala Val Cys
          50          55          60
Gln Arg Cys Gln Ala Arg Phe Ser Pro Ala Glu Arg Ile Val Asn Ser
65          70          75          80
Asn Gly Glu Leu Tyr His Glu His Cys Phe Val Cys Ala Gln Cys Phe
          85          90          95
Arg Pro Phe Pro Glu Gly Leu Phe Tyr Glu Phe Glu Gly Arg Lys Tyr
          100          105          110
Cys Glu His Asp Phe Gln Met Leu Phe Ala Pro Cys Cys Gly Ser Cys
          115          120          125
Gly Glu Phe Ile Ile Gly Arg Val Ile Lys Ala Met Asn Asn Asn Trp
          130          135          140
His Pro Gly Cys Phe Arg Cys Glu Leu Cys Asp Val Glu Leu Ala Asp
145          150          155          160
Leu Gly Phe Val Lys Asn Ala Gly Arg His Leu Cys Arg Pro Cys His
          165          170          175
Asn Arg Glu Lys Ala Lys Gly Leu Gly Lys Tyr Ile Cys Gln Arg Cys
          180          185          190
His Leu Val Ile Asp Glu Gln Pro Leu Met Phe Arg Ser Asp Ala Tyr
          195          200          205
His Pro Asp His Phe Asn Cys Thr His Cys Gly Lys Glu Leu Thr Ala
          210          215          220
Glu Ala Arg Glu Leu Lys Gly Glu Leu Tyr Cys Leu Pro Cys His Asp
225          230          235          240
Lys Met Gly Val Pro Ile Cys Gly Ala Cys Arg Arg Pro Ile Glu Gly
          245          250          255
Arg Val Val Asn Ala Leu Gly Lys Gln Trp His Val Glu His Phe Val
          260          265          270
Cys Ala Lys Cys Glu Lys Pro Phe Leu Gly His Arg His Tyr Glu Lys
          275          280          285
Lys Gly Leu Ala Tyr Cys Glu Leu
          290          295

```

<210> 682

<211> 500

<212> PRT

<213> Homo sapiens

<400> 682

```

Met Gly Ile Gly Leu Ser Ala Gln Gly Val Asn Met Asn Arg Leu Pro
 1          5          10          15
Gly Trp Asp Lys His Ser Tyr Gly Tyr His Gly Asp Asp Gly His Ser
          20          25          30
Phe Cys Ser Ser Gly Thr Gly Gln Pro Tyr Gly Pro Thr Phe Thr Thr
          35          40          45
Gly Asp Val Ile Gly Cys Cys Val Asn Leu Ile Asn Asn Thr Cys Phe
          50          55          60
Tyr Thr Lys Asn Gly His Ser Leu Gly Ile Ala Phe Thr Asp Leu Pro

```

65					70					75				80
Pro	Asn	Leu	Tyr	Pro	Thr	Val	Gly	Leu	Gln	Thr	Pro	Gly	Glu	Val
				85					90					95
Asp	Ala	Asn	Phe	Gly	Gln	His	Pro	Phe	Val	Phe	Asp	Ile	Glu	Asp
			100					105					110	Tyr
Met	Arg	Glu	Trp	Arg	Thr	Lys	Ile	Gln	Ala	Gln	Ile	Asp	Arg	Phe
		115					120					125		Pro
Ile	Gly	Asp	Arg	Glu	Gly	Glu	Trp	Gln	Thr	Met	Ile	Gln	Lys	Met
		130				135					140			Val
Ser	Ser	Tyr	Leu	Val	His	His	Gly	Tyr	Cys	Ala	Thr	Ala	Glu	Ala
145				150					155					Phe
Ala	Arg	Ser	Thr	Asp	Gln	Thr	Val	Leu	Glu	Glu	Leu	Ala	Ser	Ile
				165					170					Lys
Asn	Arg	Gln	Arg	Ile	Gln	Lys	Leu	Val	Leu	Ala	Gly	Arg	Met	Gly
		180						185					190	Glu
Ala	Ile	Glu	Thr	Thr	Gln	Gln	Leu	Tyr	Pro	Ser	Leu	Leu	Glu	Arg
	195						200					205		Asn
Pro	Asn	Leu	Leu	Phe	Thr	Leu	Lys	Val	Arg	Gln	Phe	Ile	Glu	Met
	210					215					220			Val
Asn	Gly	Thr	Asp	Ser	Glu	Val	Arg	Cys	Leu	Gly	Gly	Arg	Ser	Pro
225				230					235					Lys
Ser	Gln	Asp	Ser	Tyr	Pro	Val	Ser	Pro	Arg	Pro	Phe	Ser	Ser	Pro
			245					250						Ser
Met	Ser	Pro	Ser	His	Gly	Met	Asn	Ile	His	Asn	Leu	Ala	Ser	Gly
		260				265						270		Lys
Gly	Ser	Thr	Ala	His	Phe	Ser	Gly	Phe	Glu	Ser	Cys	Ser	Asn	Gly
	275					280					285			Val
Ile	Ser	Asn	Lys	Ala	His	Gln	Ser	Tyr	Cys	His	Ser	Asn	Lys	His
	290					295					300			Gln
Ser	Ser	Asn	Leu	Asn	Val	Pro	Glu	Leu	Asn	Ser	Ile	Asn	Met	Ser
305				310					315					Arg
Ser	Gln	Gln	Val	Asn	Asn	Phe	Thr	Ser	Asn	Asp	Val	Asp	Met	Glu
			325					330					335	Thr
Asp	His	Tyr	Ser	Asn	Gly	Val	Gly	Glu	Thr	Ser	Ser	Asn	Gly	Phe
		340					345					350		Leu
Asn	Gly	Ser	Ser	Lys	His	Asp	His	Glu	Met	Glu	Asp	Cys	Asp	Thr
	355					360					365			Glu
Met	Glu	Val	Asp	Ser	Ser	Gln	Leu	Arg	Arg	Gln	Leu	Cys	Gly	Gly
	370					375				380				Ser
Gln	Ala	Ala	Ile	Glu	Arg	Met	Ile	His	Phe	Gly	Arg	Glu	Leu	Gln
385				390					395					Ala
Met	Ser	Glu	Gln	Leu	Arg	Arg	Asp	Cys	Gly	Lys	Asn	Thr	Ala	Asn
			405					410					415	Lys
Lys	Met	Leu	Lys	Asp	Ala	Phe	Ser	Leu	Leu	Ala	Tyr	Ser	Asp	Pro
		420						425				430		Trp
Asn	Ser	Pro	Val	Gly	Asn	Gln	Leu	Asp	Pro	Ile	Gln	Arg	Glu	Pro
	435					440					445			Val
Cys	Ser	Ala	Leu	Asn	Ser	Ala	Ile	Leu	Glu	Thr	His	Asn	Leu	Pro
	450					455				460				Lys
Gln	Pro	Pro	Leu	Ala	Leu	Ala	Met	Gly	Gln	Ala	Thr	Gln	Cys	Leu
465				470					475					Gly
Leu	Met	Ala	Arg	Ser	Gly	Ile	Gly	Ser	Cys	Ala	Phe	Ala	Thr	Val
			485					490					495	Glu
Asp	Tyr	Leu	His											

500

<210> 683
 <211> 387
 <212> PRT
 <213> Homo sapiens

<400> 683

```

Met Ala Thr Ser Gly Val Leu Pro Gly Gly Gly Phe Val Ala Ser Ala
 1           5           10           15
Ala Ala Val Ala Gly Pro Glu Met Gln Thr Gly Arg Asn Asn Phe Val
 20           25           30
Ile Arg Arg Asn Pro Ala Asp Pro Gln Arg Ile Pro Ser Asn Pro Ser
 35           40           45
His Arg Ile Gln Cys Ala Ala Gly Tyr Glu Gln Ser Glu His Asn Val
 50           55           60
Cys Gln Asp Ile Asp Glu Cys Thr Ala Gly Thr His Asn Cys Arg Ala
 65           70           75           80
Asp Gln Val Cys Ile Asn Leu Arg Gly Ser Phe Ala Cys Gln Cys Pro
 85           90           95
Pro Gly Tyr Gln Lys Arg Gly Glu Gln Cys Val Asp Ile Asp Glu Cys
 100          105          110
Thr Ile Pro Pro Tyr Cys His Gln Arg Cys Val Asn Thr Pro Gly Ser
 115          120          125
Phe Tyr Cys Gln Cys Ser Pro Gly Phe Gln Leu Ala Ala Asn Asn Tyr
 130          135          140
Thr Cys Val Asp Ile Asn Glu Cys Asp Ala Ser Asn Gln Cys Ala Gln
 145          150          155          160
Gln Cys Tyr Asn Ile Leu Gly Ser Phe Ile Cys Gln Cys Asn Gln Gly
 165          170          175
Tyr Glu Leu Ser Ser Asp Arg Leu Asn Cys Glu Asp Ile Asp Glu Cys
 180          185          190
Arg Thr Ser Ser Tyr Leu Cys Gln Tyr Gln Cys Val Asn Glu Pro Gly
 195          200          205
Lys Phe Ser Cys Met Cys Pro Gln Gly Tyr Gln Val Val Arg Ser Arg
 210          215          220
Thr Cys Gln Asp Ile Asn Glu Cys Glu Thr Thr Asn Glu Cys Arg Glu
 225          230          235          240
Asp Glu Met Cys Trp Asn Tyr His Gly Gly Phe Arg Cys Tyr Pro Arg
 245          250          255
Asn Pro Cys Gln Asp Pro Tyr Ile Leu Thr Pro Glu Asn Arg Cys Val
 260          265          270
Cys Pro Val Ser Asn Ala Met Cys Arg Glu Leu Pro Gln Ser Ile Val
 275          280          285
Tyr Lys Tyr Met Ser Ile Arg Ser Asp Arg Ser Val Pro Ser Asp Ile
 290          295          300
Phe Gln Ile Gln Ala Thr Thr Ile Tyr Ala Asn Thr Ile Asn Thr Phe
 305          310          315          320
Arg Ile Lys Ser Gly Asn Glu Asn Gly Glu Phe Tyr Leu Arg Gln Thr
 325          330          335
Ser Pro Val Ser Ala Met Leu Val Leu Val Lys Ser Leu Ser Gly Pro
 340          345          350
Arg Glu His Ile Val Asp Leu Glu Met Leu Thr Val Ser Ser Ile Gly
 355          360          365

```

Thr Phe Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro
 370 375 380
 Phe Ser Phe
 385

<210> 684
 <211> 531
 <212> PRT
 <213> Homo sapiens

<400> 684
 Met Ser Lys Pro His Ser Glu Ala Gly Thr Ala Phe Ile Gln Thr Gln
 1 5 10 15
 Gln Leu His Ala Ala Met Ala Asp Thr Phe Leu Glu His Met Cys Arg
 20 25 30
 Leu Asp Ile Asp Ser Pro Pro Ile Thr Ala Arg Asn Thr Gly Ile Ile
 35 40 45
 Cys Thr Ile Gly Pro Ala Ser Arg Ser Val Glu Thr Leu Lys Glu Met
 50 55 60
 Ile Lys Ser Gly Met Asn Val Ala Arg Leu Asn Phe Ser His Gly Thr
 65 70 75 80
 His Glu Tyr His Ala Glu Thr Ile Lys Asn Val Arg Thr Ala Thr Glu
 85 90 95
 Ser Phe Ala Ser Asp Pro Tyr Leu Tyr Arg Pro Val Ala Val Ala Leu
 100 105 110
 Asp Thr Lys Gly Pro Glu Ile Arg Thr Gly Leu Ile Lys Gly Ser Gly
 115 120 125
 Thr Ala Glu Leu Glu Leu Lys Lys Gly Ala Thr Leu Lys Ile Thr Leu
 130 135 140
 Asp Asn Ala Tyr Met Glu Lys Cys Asp Glu Asn Ile Leu Trp Leu Asp
 145 150 155 160
 Tyr Lys Asn Ile Cys Lys Val Val Glu Val Gly Ser Lys Ile Tyr Val
 165 170 175
 Asp Asp Gly Leu Ile Ser Leu Gln Val Lys Gln Lys Gly Ala Asp Phe
 180 185 190
 Leu Val Thr Glu Val Glu Asn Gly Gly Ser Leu Gly Ser Lys Lys Gly
 195 200 205
 Val Asn Leu Pro Gly Ala Ala Val Asp Leu Pro Ala Val Ser Glu Lys
 210 215 220
 Asp Ile Gln Asp Leu Lys Phe Gly Val Glu Gln Asp Val Asp Met Val
 225 230 235 240
 Phe Ala Ser Phe Ile Arg Lys Ala Ser Asp Val His Glu Val Arg Lys
 245 250 255
 Val Leu Gly Glu Lys Gly Lys Asn Ile Lys Ile Ile Ser Lys Ile Glu
 260 265 270
 Asn His Glu Gly Val Arg Arg Phe Asp Glu Ile Leu Glu Ala Ser Asp
 275 280 285
 Gly Ile Met Val Ala Arg Gly Asp Leu Gly Ile Glu Ile Pro Ala Glu
 290 295 300
 Lys Val Phe Leu Ala Gln Lys Met Met Ile Gly Arg Cys Asn Arg Ala
 305 310 315 320
 Gly Lys Pro Val Ile Cys Ala Thr Gln Met Leu Glu Ser Met Ile Lys
 325 330 335
 Lys Pro Arg Pro Thr Arg Ala Glu Gly Ser Asp Val Ala Asn Ala Val

```

          340          345          350
Leu Asp Gly Ala Asp Cys Ile Met Leu Ser Gly Glu Thr Ala Lys Gly
          355          360          365
Asp Tyr Pro Leu Glu Ala Val Arg Met Gln His Leu Ile Ala Arg Glu
          370          375          380
Ala Glu Ala Ala Ile Tyr His Leu Gln Leu Phe Glu Glu Leu Arg Arg
385          390          395          400
Leu Ala Pro Ile Thr Ser Asp Pro Thr Glu Ala Thr Ala Val Gly Ala
          405          410          415
Val Glu Ala Ser Phe Lys Cys Cys Ser Gly Ala Ile Ile Val Leu Thr
          420          425          430
Lys Ser Gly Arg Ser Ala His Gln Val Ala Arg Tyr Arg Pro Arg Ala
          435          440          445
Pro Ile Ile Ala Val Thr Arg Asn Pro Gln Thr Ala Arg Gln Ala His
          450          455          460
Leu Tyr Arg Gly Ile Phe Pro Val Leu Cys Lys Asp Pro Val Gln Glu
465          470          475          480
Ala Trp Ala Glu Asp Val Asp Leu Arg Val Asn Phe Ala Met Asn Val
          485          490          495
Gly Lys Ala Arg Gly Phe Phe Lys Lys Gly Asp Val Val Ile Val Leu
          500          505          510
Thr Gly Trp Arg Pro Gly Ser Gly Phe Thr Asn Thr Met Arg Val Val
          515          520          525
Pro Val Pro
          530

```

<210> 685

<211> 480

<212> PRT

<213> Homo sapiens

<400> 685

```

Met Ala Ala Arg Cys Ser Thr Arg Trp Leu Leu Val Val Val Gly Thr
  1          5          10          15
Pro Arg Leu Pro Ala Ile Ser Gly Arg Gly Ala Arg Pro Pro Arg Glu
          20          25          30
Gly Val Val Gly Ala Trp Leu Ser Arg Lys Leu Ser Val Pro Ala Phe
          35          40          45
Ala Ser Ser Leu Thr Ser Cys Gly Pro Arg Ala Leu Leu Thr Leu Arg
          50          55          60
Pro Gly Val Ser Leu Thr Gly Thr Lys His Asn Pro Phe Ile Cys Thr
65          70          75          80
Ala Ser Phe His Thr Ser Ala Pro Leu Ala Lys Glu Asp Tyr Tyr Gln
          85          90          95
Ile Leu Gly Val Pro Arg Asn Ala Ser Gln Lys Glu Ile Lys Lys Ala
          100          105          110
Tyr Tyr Gln Leu Ala Lys Lys Tyr His Pro Asp Thr Asn Lys Asp Asp
          115          120          125
Pro Lys Ala Lys Glu Lys Phe Ser Gln Leu Ala Glu Ala Tyr Glu Val
          130          135          140
Leu Ser Asp Glu Val Lys Arg Lys Gln Tyr Asp Ala Tyr Gly Ser Ala
145          150          155          160
Gly Phe Asp Pro Gly Ala Ser Gly Ser Gln His Ser Tyr Trp Lys Gly
          165          170          175

```

Gly Pro Thr Val Asp Pro Glu Glu Leu Phe Arg Lys Ile Phe Gly Glu
 180 185 190
 Phe Ser Ser Ser Ser Phe Gly Asp Phe Gln Thr Val Phe Asp Gln Pro
 195 200 205
 Gln Glu Tyr Phe Met Glu Leu Thr Phe Asn Gln Ala Ala Lys Gly Val
 210 215 220
 Asn Lys Glu Phe Thr Val Asn Ile Met Asp Thr Cys Glu Arg Cys Asn
 225 230 235 240
 Gly Lys Gly Asn Glu Pro Gly Thr Lys Val Gln His Cys His Tyr Cys
 245 250 255
 Gly Gly Ser Gly Met Glu Thr Ile Asn Thr Gly Pro Phe Val Met Arg
 260 265 270
 Ser Thr Cys Arg Arg Cys Gly Gly Arg Gly Ser Ile Ile Ile Ser Pro
 275 280 285
 Cys Val Val Cys Arg Gly Ala Gly Gln Ala Lys Gln Lys Lys Arg Val
 290 295 300
 Met Ile Pro Val Pro Ala Gly Val Glu Asp Gly Gln Thr Val Arg Met
 305 310 315 320
 Pro Val Gly Lys Arg Glu Ile Phe Ile Thr Phe Arg Val Gln Lys Ser
 325 330 335
 Pro Val Phe Arg Arg Asp Gly Ala Asp Ile His Ser Asp Leu Phe Ile
 340 345 350
 Ser Ile Ala Gln Ala Leu Leu Gly Gly Thr Ala Arg Ala Gln Gly Leu
 355 360 365
 Tyr Glu Thr Ile Asn Val Thr Ile Pro Pro Gly Thr Gln Thr Asp Gln
 370 375 380
 Lys Ile Arg Met Gly Gly Lys Gly Ile Pro Arg Ile Asn Ser Tyr Gly
 385 390 395 400
 Tyr Gly Asp His Tyr Ile His Ile Lys Ile Arg Val Pro Lys Arg Leu
 405 410 415
 Thr Ser Arg Gln Gln Ser Leu Ile Leu Ser Tyr Ala Glu Asp Glu Thr
 420 425 430
 Asp Val Glu Gly Thr Val Asn Gly Val Thr Leu Thr Ser Ser Gly Gly
 435 440 445
 Ser Thr Met Asp Ser Ser Ala Gly Ser Lys Ala Arg Arg Glu Ala Gly
 450 455 460
 Glu Asp Glu Glu Gly Phe Leu Ser Lys Leu Lys Lys Met Phe Thr Ser
 465 470 475 480

<210> 686

<211> 572

<212> PRT

<213> Homo sapiens

<400> 686

Met Ala Ala Pro Arg Pro Ser Pro Ala Ile Ser Val Ser Val Ser Ala
 1 5 10 15
 Pro Ala Phe Tyr Ala Pro Gln Lys Lys Phe Gly Pro Val Val Ala Pro
 20 25 30
 Lys Pro Lys Val Asn Pro Phe Arg Pro Gly Asp Ser Glu Pro Pro Pro
 35 40 45
 Ala Pro Gly Ala Gln Arg Ala Gln Met Gly Arg Val Gly Glu Ile Pro
 50 55 60
 Pro Pro Pro Pro Glu Asp Phe Pro Leu Pro Pro Pro Pro Leu Ala Gly

65					70					75					80
Asp	Gly	Asp	Asp	Ala	Glu	Gly	Ala	Leu	Gly	Gly	Ala	Phe	Pro	Pro	Pro
				85					90					95	
Pro	Pro	Pro	Ile	Glu	Glu	Ser	Phe	Pro	Pro	Ala	Pro	Leu	Glu	Glu	Glu
			100					105					110		
Ile	Phe	Pro	Ser	Pro	Pro	Pro	Pro	Pro	Glu	Glu	Glu	Gly	Gly	Pro	Glu
		115					120					125			
Ala	Pro	Ile	Pro	Pro	Pro	Pro	Gln	Pro	Arg	Glu	Lys	Val	Ser	Ser	Ile
	130					135				140					
Asp	Leu	Glu	Ile	Asp	Ser	Leu	Ser	Ser	Leu	Leu	Asp	Asp	Met	Thr	Lys
145					150					155					160
Asn	Asp	Pro	Phe	Lys	Ala	Arg	Val	Ser	Ser	Gly	Tyr	Val	Pro	Pro	Pro
				165					170					175	
Val	Ala	Thr	Pro	Phe	Ser	Ser	Lys	Ser	Ser	Thr	Lys	Pro	Ala	Ala	Gly
			180					185					190		
Gly	Thr	Ala	Pro	Leu	Pro	Pro	Trp	Lys	Ser	Pro	Ser	Ser	Ser	Gln	Pro
	195						200					205			
Leu	Pro	Gln	Val	Pro	Ala	Pro	Ala	Gln	Ser	Gln	Thr	Gln	Phe	His	Val
	210					215				220					
Gln	Pro	Gln	Pro	Gln	Pro	Lys	Pro	Gln	Val	Gln	Leu	His	Val	Gln	Ser
225					230					235					240
Gln	Thr	Gln	Pro	Val	Ser	Leu	Ala	Asn	Thr	Gln	Pro	Arg	Gly	Pro	Pro
				245				250					255		
Ala	Ser	Ser	Pro	Ala	Pro	Ala	Pro	Lys	Phe	Ser	Pro	Val	Thr	Pro	Lys
			260				265					270			
Phe	Thr	Pro	Val	Ala	Ser	Lys	Phe	Ser	Pro	Gly	Ala	Pro	Gly	Gly	Ser
	275						280					285			
Gly	Ser	Gln	Pro	Asn	Gln	Lys	Leu	Gly	His	Pro	Glu	Ala	Leu	Ser	Ala
	290				295					300					
Gly	Thr	Gly	Ser	Pro	Gln	Pro	Pro	Ser	Phe	Thr	Tyr	Ala	Gln	Gln	Arg
305					310					315					320
Glu	Lys	Pro	Arg	Val	Gln	Glu	Lys	Gln	His	Pro	Val	Pro	Pro	Pro	Ala
				325				330						335	
Gln	Asn	Gln	Asn	Gln	Val	Arg	Ser	Pro	Gly	Ala	Pro	Gly	Pro	Leu	Thr
			340				345					350			
Leu	Lys	Glu	Val	Glu	Glu	Leu	Glu	Gln	Leu	Thr	Gln	Gln	Leu	Met	Gln
	355					360					365				
Asp	Met	Glu	His	Pro	Gln	Arg	Gln	Asn	Val	Ala	Val	Asn	Glu	Leu	Cys
	370				375					380					
Gly	Arg	Cys	His	Gln	Pro	Leu	Ala	Arg	Ala	Gln	Pro	Ala	Val	Arg	Ala
385					390					395					400
Leu	Gly	Gln	Leu	Phe	His	Ile	Ala	Cys	Phe	Thr	Cys	His	Gln	Cys	Ala
				405				410					415		
Gln	Gln	Leu	Gln	Gly	Gln	Gln	Phe	Tyr	Ser	Leu	Glu	Gly	Ala	Pro	Tyr
			420				425					430			
Cys	Glu	Gly	Cys	Tyr	Thr	Asp	Thr	Leu	Glu	Lys	Cys	Asn	Thr	Cys	Gly
	435					440					445				
Glu	Pro	Ile	Thr	Asp	Arg	Met	Leu	Arg	Ala	Thr	Gly	Lys	Ala	Tyr	His
	450				455					460					
Pro	His	Cys	Phe	Thr	Cys	Val	Val	Cys	Ala	Arg	Pro	Leu	Glu	Gly	Thr
465					470					475					480
Ser	Phe	Ile	Val	Asp	Gln	Ala	Asn	Arg	Pro	His	Cys	Val	Pro	Asp	Tyr
				485				490					495		
His	Lys	Gln	Tyr	Ala	Pro	Arg	Cys	Ser	Val	Cys	Ser	Glu	Pro	Ile	Met

			500					505					510				
Pro	Glu	Pro	Gly	Arg	Asp	Glu	Thr	Val	Arg	Val	Val	Ala	Leu	Asp	Lys		
		515						520				525					
Asn	Phe	His	Met	Lys	Cys	Tyr	Lys	Cys	Glu	Asp	Cys	Gly	Lys	Pro	Leu		
	530					535					540						
Ser	Ile	Glu	Ala	Asp	Asp	Asn	Gly	Cys	Phe	Pro	Leu	Asp	Gly	His	Val		
545					550					555					560		
Leu	Cys	Arg	Lys	Cys	His	Thr	Ala	Arg	Ala	Gln	Thr						
			565					570									

<210> 687

<211> 2861

<212> PRT

<213> Homo sapiens

<400> 687

Met	Lys	Ala	Met	Asp	Val	Leu	Pro	Ile	Leu	Lys	Glu	Lys	Val	Ala	Tyr		
1			5					10						15			
Leu	Ser	Gly	Gly	Arg	Asp	Lys	Arg	Gly	Gly	Pro	Ile	Leu	Thr	Phe	Pro		
		20						25					30				
Ala	Arg	Ser	Asn	His	Asp	Arg	Ile	Arg	Gln	Glu	Asp	Leu	Arg	Arg	Leu		
	35					40						45					
Ile	Ser	Tyr	Leu	Ala	Cys	Ile	Pro	Ser	Glu	Glu	Val	Cys	Lys	Arg	Gly		
	50				55						60						
Phe	Thr	Val	Ile	Val	Asp	Met	Arg	Gly	Ser	Lys	Trp	Asp	Ser	Ile	Lys		
65					70					75					80		
Pro	Leu	Leu	Lys	Ile	Leu	Gln	Glu	Ser	Phe	Pro	Cys	Cys	Ile	His	Val		
			85					90						95			
Ala	Leu	Ile	Ile	Lys	Pro	Asp	Asn	Phe	Trp	Gln	Lys	Gln	Arg	Thr	Asn		
		100					105						110				
Phe	Gly	Ser	Ser	Lys	Phe	Glu	Phe	Glu	Thr	Asn	Met	Val	Ser	Leu	Glu		
	115					120						125					
Gly	Leu	Thr	Lys	Val	Val	Asp	Pro	Ser	Gln	Leu	Thr	Pro	Glu	Phe	Asp		
	130					135					140						
Gly	Cys	Leu	Glu	Tyr	Asn	His	Glu	Glu	Trp	Ile	Glu	Ile	Arg	Val	Ala		
145				150					155						160		
Phe	Glu	Asp	Tyr	Ile	Ser	Asn	Ala	Thr	His	Met	Leu	Ser	Arg	Leu	Glu		
			165					170						175			
Glu	Leu	Gln	Asp	Ile	Leu	Ala	Lys	Lys	Glu	Leu	Pro	Gln	Asp	Leu	Glu		
		180						185					190				
Gly	Ala	Arg	Asn	Met	Ile	Glu	Glu	His	Ser	Gln	Leu	Lys	Lys	Lys	Val		
	195					200						205					
Ile	Lys	Ala	Pro	Ile	Glu	Asp	Leu	Asp	Leu	Glu	Gly	Gln	Lys	Leu	Leu		
	210				215						220						
Gln	Arg	Ile	Gln	Ser	Ser	Glu	Ser	Phe	Pro	Lys	Lys	Asn	Ser	Gly	Ser		
225				230						235					240		
Gly	Asn	Ala	Asp	Leu	Gln	Asn	Leu	Leu	Pro	Lys	Val	Ser	Thr	Met	Leu		
		245						250						255			
Asp	Arg	Leu	His	Ser	Thr	Arg	Gln	His	Leu	His	Gln	Met	Trp	His	Val		
		260					265						270				
Arg	Lys	Leu	Lys	Leu	Asp	Gln	Cys	Phe	Gln	Leu	Arg	Leu	Phe	Glu	Gln		
	275					280						285					
Asp	Ala	Glu	Lys	Met	Phe	Asp	Trp	Ile	Thr	His	Asn	Lys	Gly	Leu	Phe		
	290					295					300						

Leu	Asn	Ser	Tyr	Thr	Glu	Ile	Gly	Thr	Ser	His	Pro	His	Ala	Met	Glu
305					310					315					320
Leu	Gln	Thr	Gln	His	Asn	His	Phe	Ala	Met	Asn	Cys	Met	Asn	Val	Tyr
				325					330					335	
Val	Asn	Ile	Asn	Arg	Ile	Met	Ser	Val	Ala	Asn	Arg	Leu	Val	Glu	Ser
			340					345					350		
Gly	His	Tyr	Ala	Ser	Gln	Gln	Ile	Arg	Gln	Ile	Ala	Ser	Gln	Leu	Glu
	355					360					365				
Gln	Glu	Trp	Lys	Ala	Phe	Ala	Ala	Ala	Leu	Asp	Glu	Arg	Ser	Thr	Leu
370					375						380				
Leu	Asp	Met	Ser	Ser	Ile	Phe	His	Gln	Lys	Ala	Glu	Lys	Tyr	Met	Ser
385					390					395					400
Asn	Val	Asp	Ser	Trp	Cys	Lys	Ala	Cys	Gly	Glu	Val	Asp	Leu	Pro	Ser
			405						410					415	
Glu	Leu	Gln	Asp	Leu	Glu	Asp	Ala	Ile	His	His	His	Gln	Gly	Ile	Tyr
			420					425					430		
Glu	His	Ile	Thr	Leu	Ala	Tyr	Ser	Glu	Val	Ser	Gln	Asp	Gly	Lys	Ser
	435						440				445				
Leu	Leu	Asp	Lys	Leu	Gln	Arg	Pro	Leu	Thr	Pro	Gly	Ser	Ser	Asp	Ser
450					455						460				
Leu	Thr	Ala	Ser	Ala	Asn	Tyr	Ser	Lys	Ala	Val	His	His	Val	Leu	Asp
465					470					475					480
Val	Ile	His	Glu	Val	Leu	His	His	Gln	Arg	His	Val	Arg	Thr	Ile	Trp
			485					490						495	
Gln	His	Arg	Lys	Val	Arg	Leu	His	Gln	Arg	Leu	Gln	Leu	Cys	Val	Phe
			500					505					510		
Gln	Gln	Glu	Val	Gln	Gln	Val	Leu	Asp	Trp	Ile	Glu	Asn	His	Gly	Glu
	515						520				525				
Ala	Phe	Leu	Ser	Lys	His	Thr	Gly	Val	Gly	Lys	Ser	Leu	His	Arg	Ala
530					535						540				
Arg	Ala	Leu	Gln	Lys	Arg	His	Glu	Asp	Phe	Glu	Glu	Val	Ala	Gln	Asn
545				550						555					560
Thr	Tyr	Thr	Asn	Ala	Asp	Lys	Leu	Leu	Glu	Ala	Ala	Glu	Gln	Leu	Ala
			565					570						575	
Gln	Thr	Gly	Glu	Cys	Asp	Pro	Glu	Glu	Ile	Tyr	Gln	Ala	Ala	His	Gln
		580					585						590		
Leu	Glu	Asp	Arg	Ile	Gln	Asp	Phe	Val	Arg	Arg	Val	Glu	Gln	Arg	Lys
	595					600					605				
Ile	Leu	Leu	Asp	Met	Ser	Val	Ser	Phe	His	Thr	His	Val	Lys	Glu	Leu
610					615					620					
Trp	Thr	Trp	Leu	Glu	Glu	Leu	Gln	Lys	Glu	Leu	Leu	Asp	Asp	Val	Tyr
625				630					635						640
Ala	Glu	Ser	Val	Glu	Ala	Val	Gln	Asp	Leu	Ile	Lys	Arg	Phe	Gly	Gln
			645				650						655		
Gln	Gln	Gln	Thr	Thr	Leu	Gln	Val	Thr	Val	Asn	Val	Ile	Lys	Glu	Gly
		660					665					670			
Glu	Asp	Leu	Ile	Gln	Gln	Leu	Arg	Asp	Ser	Ala	Ile	Ser	Ser	Asn	Lys
	675					680					685				
Thr	Pro	His	Asn	Ser	Ser	Ile	Asn	His	Ile	Glu	Thr	Val	Leu	Gln	Gln
690					695					700					
Leu	Asp	Glu	Ala	Gln	Ser	Gln	Met	Glu	Glu	Leu	Phe	Gln	Glu	Arg	Lys
705				710					715						720
Ile	Lys	Leu	Glu	Leu	Phe	Leu	His	Val	Arg	Ile	Phe	Glu	Arg	Asp	Ala
			725					730						735	

-254-

Tyr	Arg	Asp	Phe	Ser	Leu	Arg	Met	Glu	Lys	Tyr	Arg	Thr	Ser	Leu	Glu	1170	1175	1180
Lys	Ala	Leu	Gly	Ile	Ser	Ser	Asp	Ser	Asn	Lys	Ser	Ser	Lys	Ser	Leu	1185	1190	1195
Gln	Leu	Asp	Ile	Ile	Pro	Ala	Ser	Ile	Pro	Gly	Ser	Glu	Val	Lys	Leu	1205	1210	1215
Arg	Asp	Ala	Ala	His	Glu	Leu	Asn	Glu	Glu	Lys	Arg	Lys	Ser	Ala	Arg	1220	1225	1230
Arg	Lys	Glu	Phe	Ile	Met	Ala	Glu	Leu	Ile	Gln	Thr	Glu	Lys	Ala	Tyr	1235	1240	1245
Val	Arg	Asp	Leu	Arg	Glu	Cys	Met	Asp	Thr	Tyr	Leu	Trp	Glu	Met	Thr	1250	1255	1260
Ser	Gly	Val	Glu	Glu	Ile	Pro	Pro	Gly	Ile	Val	Asn	Lys	Glu	Leu	Ile	1265	1270	1275
Ile	Phe	Gly	Asn	Met	Gln	Glu	Ile	Tyr	Glu	Phe	His	Asn	Asn	Ile	Phe	1285	1290	1295
Leu	Lys	Glu	Leu	Glu	Lys	Tyr	Glu	Gln	Leu	Pro	Glu	Asp	Val	Gly	His	1300	1305	1310
Cys	Phe	Val	Thr	Trp	Ala	Asp	Lys	Phe	Gln	Met	Tyr	Val	Thr	Tyr	Cys	1315	1320	1325
Lys	Asn	Lys	Pro	Asp	Ser	Thr	Gln	Leu	Ile	Leu	Glu	His	Ala	Gly	Ser	1330	1335	1340
Tyr	Phe	Asp	Glu	Ile	Gln	Gln	Arg	His	Gly	Leu	Ala	Asn	Ser	Ile	Ser	1345	1350	1355
Ser	Tyr	Leu	Ile	Lys	Pro	Val	Gln	Arg	Ile	Thr	Lys	Tyr	Gln	Leu	Leu	1365	1370	1375
Leu	Lys	Glu	Leu	Leu	Thr	Cys	Cys	Glu	Glu	Gly	Lys	Gly	Glu	Ile	Lys	1380	1385	1390
Asp	Gly	Leu	Glu	Val	Met	Leu	Ser	Val	Pro	Lys	Arg	Ala	Asn	Asp	Ala	1395	1400	1405
Met	His	Leu	Ser	Met	Leu	Glu	Gly	Phe	Asp	Glu	Asn	Ile	Glu	Ser	Gln	1410	1415	1420
Gly	Glu	Leu	Ile	Leu	Gln	Glu	Ser	Phe	Gln	Val	Trp	Asp	Pro	Lys	Thr	1425	1430	1435
Leu	Ile	Arg	Lys	Gly	Arg	Glu	Arg	His	Leu	Phe	Leu	Phe	Glu	Met	Ser	1445	1450	1455
Leu	Val	Phe	Ser	Lys	Glu	Val	Lys	Asp	Ser	Ser	Gly	Arg	Ser	Lys	Tyr	1460	1465	1470
Leu	Tyr	Lys	Ser	Lys	Leu	Phe	Thr	Ser	Glu	Leu	Gly	Val	Thr	Glu	His	1475	1480	1485
Val	Glu	Gly	Asp	Pro	Cys	Lys	Phe	Ala	Leu	Trp	Val	Gly	Arg	Thr	Pro	1490	1495	1500
Thr	Ser	Asp	Asn	Lys	Ile	Val	Leu	Lys	Ala	Ser	Ser	Ile	Glu	Asn	Lys	1505	1510	1515
Gln	Asp	Trp	Ile	Lys	His	Ile	Arg	Glu	Val	Ile	Gln	Glu	Arg	Thr	Ile	1525	1530	1535
His	Leu	Lys	Gly	Ala	Leu	Lys	Glu	Pro	Ile	His	Ile	Pro	Lys	Thr	Ala	1540	1545	1550
Pro	Ala	Thr	Arg	Gln	Lys	Gly	Arg	Arg	Asp	Gly	Glu	Asp	Leu	Asp	Ser	1555	1560	1565
Gln	Gly	Asp	Gly	Ser	Ser	Gln	Pro	Asp	Thr	Ile	Ser	Ile	Ala	Ser	Arg	1570	1575	1580
Thr	Ser	Gln	Asn	Thr	Leu	Asp	Ser	Asp	Lys	Leu	Ser	Gly	Gly	Cys	Glu	1585	1590	1595

Leu Thr Val Val Ile His Asp Phe Thr Ala Cys Asn Ser Asn Glu Leu
 1605 1610 1615
 Thr Ile Arg Arg Gly Gln Thr Val Glu Val Leu Glu Arg Pro His Asp
 1620 1625 1630
 Lys Pro Asp Trp Cys Leu Val Arg Thr Thr Asp Arg Ser Pro Ala Ala
 1635 1640 1645
 Glu Gly Leu Val Pro Cys Gly Ser Leu Cys Ile Ala His Ser Arg Ser
 1650 1655 1660
 Ser Met Glu Met Glu Gly Ile Phe Asn His Lys Asp Ser Leu Ser Val
 1665 1670 1675 1680
 Ser Ser Asn Asp Ala Ser Pro Pro Ala Ser Val Ala Ser Leu Gln Pro
 1685 1690 1695
 His Met Ile Gly Ala Gln Ser Ser Pro Gly Pro Lys Arg Pro Gly Asn
 1700 1705 1710
 Thr Leu Arg Lys Trp Leu Thr Ser Pro Val Arg Arg Leu Ser Ser Gly
 1715 1720 1725
 Lys Ala Asp Gly His Val Lys Lys Leu Ala His Lys His Lys Lys Ser
 1730 1735 1740
 Arg Glu Val Arg Lys Ser Ala Asp Ala Gly Ser Gln Lys Asp Ser Asp
 1745 1750 1755 1760
 Asp Ser Ala Ala Thr Pro Gln Asp Glu Thr Val Glu Glu Arg Gly Arg
 1765 1770 1775
 Asn Glu Gly Leu Ser Ser Gly Thr Leu Ser Lys Ser Ser Ser Ser Gly
 1780 1785 1790
 Met Gln Ser Cys Gly Glu Glu Glu Gly Glu Glu Gly Ala Asp Ala Val
 1795 1800 1805
 Pro Leu Pro Pro Pro Met Ala Ile Gln Gln His Ser Leu Leu Gln Pro
 1810 1815 1820
 Asp Ser Gln Asp Asp Lys Ala Ser Ser Arg Leu Leu Val Arg Pro Thr
 1825 1830 1835 1840
 Ser Ser Glu Thr Pro Ser Ala Ala Glu Leu Val Ser Ala Ile Glu Glu
 1845 1850 1855
 Leu Val Lys Ser Lys Met Ala Leu Glu Asp Arg Pro Ser Ser Leu Leu
 1860 1865 1870
 Val Asp Gln Gly Asp Ser Ser Ser Pro Ser Phe Asn Pro Ser Asp Asn
 1875 1880 1885
 Ser Leu Leu Ser Ser Ser Ser Pro Ile Asp Glu Met Glu Glu Arg Lys
 1890 1895 1900
 Ser Ser Ser Leu Lys Arg Arg His Tyr Val Leu Gln Glu Leu Val Glu
 1905 1910 1915 1920
 Thr Glu Arg Asp Tyr Val Arg Asp Leu Gly Tyr Val Val Glu Gly Tyr
 1925 1930 1935
 Met Ala Leu Met Lys Glu Asp Gly Val Pro Asp Asp Met Lys Gly Lys
 1940 1945 1950
 Asp Lys Ile Val Phe Gly Asn Ile His Gln Ile Tyr Asp Trp His Arg
 1955 1960 1965
 Asp Phe Phe Leu Gly Glu Leu Glu Lys Cys Leu Glu Asp Pro Glu Lys
 1970 1975 1980
 Leu Gly Ser Leu Phe Val Lys His Glu Arg Arg Leu His Met Tyr Ile
 1985 1990 1995 2000
 Ala Tyr Cys Gln Asn Lys Pro Lys Ser Glu His Ile Val Ser Glu Tyr
 2005 2010 2015
 Ile Asp Thr Phe Glu Asp Leu Lys Gln Arg Leu Gly His Arg Leu
 2020 2025 2030

Gln Leu Thr Asp Leu Leu Ile Lys Pro Val Gln Arg Ile Met Lys Tyr
 2035 2040 2045
 Gln Leu Leu Leu Lys Asp Phe Leu Lys Tyr Ser Lys Lys Ala Ser Leu
 2050 2055 2060
 Asp Thr Ser Glu Leu Glu Arg Ala Val Glu Val Met Cys Ile Val Pro
 2065 2070 2075 2080
 Arg Arg Cys Asn Asp Met Met Asn Val Gly Arg Leu Gln Gly Phe Asp
 2085 2090 2095
 Gly Lys Ile Val Ala Gln Gly Lys Leu Leu Leu Gln Asp Thr Phe Leu
 2100 2105 2110
 Val Thr Asp Gln Asp Ala Gly Leu Leu Pro Arg Cys Arg Glu Arg Arg
 2115 2120 2125
 Ile Phe Leu Phe Glu Gln Ile Val Ile Phe Ser Glu Pro Leu Asp Lys
 2130 2135 2140
 Lys Lys Gly Phe Ser Met Pro Gly Phe Leu Phe Lys Asn Ser Ile Lys
 2145 2150 2155 2160
 Val Ser Cys Leu Cys Leu Glu Glu Asn Val Glu Asn Asp Pro Cys Lys
 2165 2170 2175
 Phe Ala Leu Thr Ser Arg Thr Gly Asp Val Val Glu Thr Phe Ile Leu
 2180 2185 2190
 His Ser Ser Ser Pro Ser Val Arg Gln Thr Trp Ile His Glu Ile Asn
 2195 2200 2205
 Gln Ile Leu Glu Asn Gln Arg Asn Phe Leu Asn Ala Leu Thr Ser Pro
 2210 2215 2220
 Ile Glu Tyr Gln Arg Asn His Ser Gly Gly Gly Gly Gly Gly Gly Ser
 2225 2230 2235 2240
 Gly Ala Ala Ala Gly Val Gly Ala Ala Ala Ala Ala Gly Pro Pro Val
 2245 2250 2255
 Ala Ala Ala Ala Thr Val Ala Ala Pro Ala Ala Ala Ala Ala Pro Pro
 2260 2265 2270
 Ala Arg Ala Gly Ala Gly Pro Pro Gly Ser Pro Ser Leu Ser Asp Thr
 2275 2280 2285
 Thr Pro Pro Cys Trp Ser Pro Leu Gln Pro Arg Ala Arg Gln Arg Gln
 2290 2295 2300
 Thr Arg Cys Gln Ser Glu Ser Ser Ser Ser Ser Asn Ile Ser Thr Met
 2305 2310 2315 2320
 Leu Val Thr His Asp Tyr Thr Ala Val Lys Glu Asp Glu Ile Asn Val
 2325 2330 2335
 Tyr Gln Gly Glu Val Val Gln Ile Leu Ala Ser Asn Gln Gln Asn Met
 2340 2345 2350
 Phe Leu Val Phe Arg Ala Ala Thr Asp Gln Cys Pro Ala Ala Glu Gly
 2355 2360 2365
 Trp Ile Pro Gly Phe Val Leu Gly His Thr Ser Ala Val Ile Val Glu
 2370 2375 2380
 Asn Pro Asp Gly Thr Leu Lys Lys Ser Thr Ser Trp His Thr Ala Leu
 2385 2390 2395 2400
 Arg Leu Arg Lys Lys Ser Glu Lys Lys Asp Lys Asp Gly Lys Arg Glu
 2405 2410 2415
 Gly Lys Leu Glu Asn Gly Tyr Arg Lys Ser Arg Glu Gly Leu Ser Asn
 2420 2425 2430
 Lys Val Ser Val Lys Leu Leu Asn Pro Asn Tyr Ile Tyr Asp Val Pro
 2435 2440 2445
 Pro Glu Phe Val Ile Pro Leu Ser Glu Val Thr Cys Glu Thr Gly Glu
 2450 2455 2460

Thr Val Val Leu Arg Cys Arg Val Cys Gly Arg Pro Lys Ala Ser Ile
 2465 2470 2475 2480
 Thr Trp Lys Gly Pro Glu His Asn Thr Leu Asn Asn Asp Gly His Tyr
 2485 2490 2495
 Ser Ile Ser Tyr Ser Asp Leu Gly Glu Ala Thr Leu Lys Ile Val Gly
 2500 2505 2510
 Val Thr Thr Glu Asp Asp Gly Ile Tyr Thr Cys Ile Ala Val Asn Asp
 2515 2520 2525
 Met Gly Ser Ala Ser Ser Ser Ala Ser Leu Arg Val Leu Gly Pro Gly
 2530 2535 2540
 Met Asp Gly Ile Met Val Thr Trp Lys Asp Asn Phe Asp Ser Phe Tyr
 2545 2550 2555 2560
 Ser Glu Val Ala Glu Leu Gly Arg Gly Arg Phe Ser Val Val Lys Lys
 2565 2570 2575
 Cys Asp Gln Lys Gly Thr Lys Arg Ala Val Ala Thr Lys Phe Val Asn
 2580 2585 2590
 Lys Lys Leu Met Lys Arg Asp Gln Val Thr His Glu Leu Gly Ile Leu
 2595 2600 2605
 Gln Ser Leu Gln His Pro Leu Leu Val Gly Leu Leu Asp Thr Phe Glu
 2610 2615 2620
 Thr Pro Thr Ser Tyr Ile Leu Val Leu Glu Met Ala Asp Gln Gly Arg
 2625 2630 2635 2640
 Leu Leu Asp Cys Val Val Arg Trp Gly Ser Leu Thr Glu Gly Lys Ile
 2645 2650 2655
 Arg Ala His Leu Gly Glu Val Leu Glu Ala Val Arg Tyr Leu His Asn
 2660 2665 2670
 Cys Arg Ile Ala His Leu Asp Leu Lys Pro Glu Asn Ile Leu Val Asp
 2675 2680 2685
 Glu Ser Leu Ala Lys Pro Thr Ile Lys Leu Ala Asp Phe Gly Asp Ala
 2690 2695 2700
 Val Gln Leu Asn Thr Thr Tyr Tyr Ile His Gln Leu Leu Gly Asn Pro
 2705 2710 2715 2720
 Glu Phe Ala Ala Pro Glu Ile Ile Leu Gly Asn Pro Val Ser Leu Thr
 2725 2730 2735
 Ser Asp Thr Trp Ser Val Gly Val Leu Thr Tyr Val Leu Leu Ser Gly
 2740 2745 2750
 Val Ser Pro Phe Leu Asp Asp Ser Val Glu Glu Thr Cys Leu Asn Ile
 2755 2760 2765
 Cys Arg Leu Asp Phe Ser Phe Pro Asp Asp Tyr Phe Lys Gly Val Ser
 2770 2775 2780
 Gln Lys Ala Lys Glu Phe Val Cys Phe Leu Leu Gln Glu Asp Pro Ala
 2785 2790 2795 2800
 Lys Arg Pro Ser Ala Ala Leu Ala Leu Gln Glu Gln Trp Leu Gln Ala
 2805 2810 2815
 Gly Asn Gly Arg Ser Thr Gly Val Leu Asp Thr Ser Arg Leu Thr Ser
 2820 2825 2830
 Phe Ile Glu Arg Arg Lys His Gln Asn Asp Val Arg Pro Ile Arg Ser
 2835 2840 2845
 Ile Lys Asn Phe Leu Gln Ser Arg Leu Leu Pro Arg Val
 2850 2855 2860

<210> 688

<211> 271

<212> PRT

<213> Homo sapiens

<400> 688

```

Met Val Leu Ile Lys Glu Phe Arg Val Val Leu Pro Cys Ser Val Gln
 1           5           10           15
Glu Tyr Gln Val Gly Gln Leu Tyr Ser Val Ala Glu Ala Ser Lys Asn
          20           25           30
Glu Thr Gly Gly Gly Glu Gly Ile Glu Val Leu Lys Asn Glu Pro Tyr
          35           40           45
Glu Lys Asp Gly Glu Lys Gly Gln Tyr Thr His Lys Ile Tyr His Leu
          50           55           60
Lys Ser Lys Val Pro Ala Phe Val Arg Met Ile Ala Pro Glu Gly Ser
65           70           75           80
Leu Val Phe His Glu Lys Ala Trp Asn Ala Tyr Pro Tyr Cys Arg Thr
          85           90           95
Ile Val Thr Asn Glu Tyr Met Lys Asp Asp Phe Phe Ile Lys Ile Glu
          100          105          110
Thr Trp His Lys Pro Asp Leu Gly Thr Leu Glu Asn Val His Gly Leu
          115          120          125
Asp Pro Asn Thr Trp Lys Thr Val Glu Ile Val His Ile Asp Ile Ala
          130          135          140
Asp Arg Ser Gln Val Glu Pro Ala Asp Tyr Lys Ala Asp Glu Asp Pro
145          150          155          160
Ala Leu Phe Gln Ser Val Lys Thr Lys Arg Gly Pro Leu Gly Pro Asn
          165          170          175
Trp Lys Lys Glu Leu Ala Asn Ser Pro Asp Cys Pro Gln Met Cys Ala
          180          185          190
Tyr Lys Leu Val Thr Ile Lys Phe Lys Trp Trp Gly Leu Gln Ser Lys
          195          200          205
Val Glu Asn Phe Ile Gln Lys Gln Glu Lys Arg Ile Phe Thr Asn Phe
          210          215          220
His Arg Gln Leu Phe Cys Trp Ile Asp Lys Trp Ile Asp Leu Thr Met
225          230          235          240
Glu Asp Ile Arg Arg Met Glu Asp Glu Thr Gln Lys Glu Leu Glu Thr
          245          250          255
Met Arg Lys Arg Gly Ser Val Arg Gly Thr Ser Ala Ala Asp Val
          260          265          270

```

<210> 642

<211> 11

<212> DNA

<213> Homo sapiens

<400> 642

gatgatgatg a

11

<210> 643

<211> 11

<212> DNA

<213> Homo sapiens

<400> 643

gatgatgatg a

11

<210> 644
<211> 11
<212> DNA
<213> Homo sapiens

<400> 644
gatgatgatg a 11

<210> 645
<211> 11
<212> DNA
<213> Homo sapiens

<400> 645
gatgatgatg a 11

<210> 646
<211> 11
<212> DNA
<213> Homo sapiens

<400> 646
gatgatgatg a 11

<210> 647
<211> 11
<212> DNA
<213> Homo sapiens

<400> 647
gatgatgatg a 11

<210> 648
<211> 11
<212> DNA
<213> Homo sapiens

<400> 648
gatgatgatg a 11

<210> 649
<211> 11
<212> DNA
<213> Homo sapiens

<400> 649
gatgatgatg a 11

<210> 650
<211> 11
<212> DNA
<213> Homo sapiens

<400> 650

gatgatgatg a	11
<210> 651	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 651	
gatgatgatg a	11
<210> 652	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 652	
gatgatgatg a	11
<210> 653	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 653	
gatgatgatg a	11
<210> 654	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 654	
gatgatgatg a	11
<210> 655	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 655	
gatgatgatg a	11
<210> 656	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 656	
gatgatgatg a	11
<210> 657	
<211> 11	
<212> DNA	
<213> Homo sapiens	

<400> 657 gatgatgatg a	11
<210> 658 <211> 11 <212> DNA <213> Homo sapiens	
<400> 658 gatgatgatg a	11
<210> 659 <211> 11 <212> DNA <213> Homo sapiens	
<400> 659 gatgatgatg a	11
<210> 660 <211> 11 <212> DNA <213> Homo sapiens	
<400> 660 gatgatgatg a	11
<210> 661 <211> 11 <212> DNA <213> Homo sapiens	
<400> 661 gatgatgatg a	11
<210> 662 <211> 11 <212> DNA <213> Homo sapiens	
<400> 662 gatgatgatg a	11
<210> 663 <211> 11 <212> DNA <213> Homo sapiens	
<400> 663 gatgatgatg a	11
<210> 664 <211> 11 <212> DNA	

<213> Homo sapiens

<400> 664
gatgatgatg a

11

<210> 665
<211> 11
<212> DNA
<213> Homo sapiens

<400> 665
gatgatgatg a

11

<210> 666
<211> 11
<212> DNA
<213> Homo sapiens

<400> 666
gatgatgatg a

11

<210> 667
<211> 11
<212> DNA
<213> Homo sapiens

<400> 667
gatgatgatg a

11

<210> 668
<211> 11
<212> DNA
<213> Homo sapiens

<400> 668
gatgatgatg a

11

<210> 669
<211> 11
<212> DNA
<213> Homo sapiens

<400> 669
gatgatgatg a

11

<210> 670
<211> 11
<212> DNA
<213> Homo sapiens

<400> 670
gatgatgatg a

11

<210> 671

<211> 11
<212> DNA
<213> Homo sapiens

<400> 671
gatgatgatg a 11

<210> 672
<211> 11
<212> DNA
<213> Homo sapiens

<400> 672
gatgatgatg a 11

<210> 673
<211> 11
<212> DNA
<213> Homo sapiens

<400> 673
gatgatgatg a 11

<210> 674
<211> 11
<212> DNA
<213> Homo sapiens

<400> 674
gatgatgatg a 11

<210> 675
<211> 11
<212> DNA
<213> Homo sapiens

<400> 675
gatgatgatg a 11

<210> 676
<211> 11
<212> DNA
<213> Homo sapiens

<400> 676
gatgatgatg a 11

<210> 677
<211> 11
<212> DNA
<213> Homo sapiens

<400> 677
gatgatgatg a 11

<210> 678
<211> 11
<212> DNA
<213> Homo sapiens

<400> 678
gatgatgatg a 11

<210> 679
<211> 11
<212> DNA
<213> Homo sapiens

<400> 679
gatgatgatg a 11

<210> 680
<211> 11
<212> DNA
<213> Homo sapiens

<400> 680
gatgatgatg a 11

<210> 681
<211> 11
<212> DNA
<213> Homo sapiens

<400> 681
gatgatgatg a 11

<210> 682
<211> 11
<212> DNA
<213> Homo sapiens

<400> 682
gatgatgatg a 11

<210> 683
<211> 11
<212> DNA
<213> Homo sapiens

<400> 683
gatgatgatg a 11

<210> 684
<211> 11
<212> DNA
<213> Homo sapiens

<400> 684

gatgatgatg a	11
<210> 685	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 685	
gatgatgatg a	11
<210> 686	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 686	
gatgatgatg a	11
<210> 687	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 687	
gatgatgatg a	11
<210> 688	
<211> 11	
<212> DNA	
<213> Homo sapiens	
<400> 688	
gatgatgatg a	11
<210> 689	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 689	
gtcagcctgg aggagttctc a	21
<210> 690	
<211> 22	
<212> DNA	
<213> Homo sapiens	
<400> 690	
tcacccttgg caatacagat gt	22
<210> 691	
<211> 21	
<212> DNA	
<213> Homo sapiens	

<400> 691
cccacccatg tgcccgtgac a 21

<210> 692
<211> 19
<212> DNA
<213> Homo sapiens

<400> 692
cgtgattgcc gacgatctc 19

<210> 693
<211> 19
<212> DNA
<213> Homo sapiens

<400> 693
ttccggccgc tagtcttgt 19

<210> 694
<211> 26
<212> DNA
<213> Homo sapiens

<400> 694
cgcaccggtt cgtctgacg cagtac 26

<210> 695
<211> 22
<212> DNA
<213> Homo sapiens

<400> 695
ctttcccccac gagtatgttg gt 22

<210> 696
<211> 19
<212> DNA
<213> Homo sapiens

<400> 696
aagggaccgt gctgtgagc 19

<210> 697
<211> 29
<212> DNA
<213> Homo sapiens

<400> 697
agccctcat gtgcctctca acttcatag 29

<210> 698
<211> 6
<212> PRT

<213> Artificial Sequence

<220>

<223> partial signal sequence of HBM and Zmax1

<400> 698

Ser Pro Ala His Ser Ser

1

5

<210> 699

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic duplex

<400> 699

tcgaccacca tggaggcagc gccgc

25

<210> 700

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic duplex

<400> 700

ggtggtacct ccgtcgcggc gggcc

25

<210> 701

<211> 55

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic duplex

<400> 701

catgtgtgaa atcaccaagc cgccctcaga cgacagcccg gccacagca gtggc

55

<210> 702

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic duplex

<400> 702

tcgagtacac acttttagtgg ttcggcgggg gtctgtgtgc gggccgggtg tcgtcaccgc 60

cgg

63

<210> 703
 <211> 17
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 703
 Val Glu Thr Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg
 1 5 10 15
 Lys

<210> 704
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 704
 Asn Lys Arg Thr Gly Gly Lys Arg Lys Glu Ile Leu Ser Ala Leu
 1 5 10 15

<210> 705
 <211> 12
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 705
 Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
 1 5 10

<210> 706
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 706
 Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys Leu Pro
 1 5 10 15

<210> 707
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 707
Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln Cys Gly Gln
1 5 10 15

<210> 708
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 708
Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
1 5 10

<210> 709
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 709
Tyr Ala Gly Ala Asn Gly Pro Phe Pro His Glu Tyr Val Ser Gly
1 5 10 15

<210> 710
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 710
Ala Cys Gly Lys Ser Met Met Ser Ser Val Ser Leu Met
1 5 10

<210> 711
<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 711

Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser Asp Ser Asp

1

5

10

15

Pro Tyr

<210> 712

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 712

Ser Gly Trp Asn Glu Cys Ala Ser Ser Asn Gly His Cys Ser His

1

5

10

15

<210> 713

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide from Zmax1 sequence (SEQ ID NO: 3)

<400> 713

Asn Gly Asp Ala Asn Cys Gln Asp Lys Ser Asp Glu Lys Asn

1

5

10

<210> 714

<211> 22

<212> PRT

<213> Homo sapiens

<400> 714

Ile Lys Arg Ala Lys Asp Asp Gly Thr Gln Pro Phe Val Leu Thr Ser

1

5

10

15

Leu Ser Gln Gly Gln Asn

20

<210> 715

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 715

cgtacgtaaa gcggccgctt ggcaatacag atgtggga

38

<210> 716

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 716

cgtacgtaaa gcggccgcag tagctcctct cggtggc

37

<210> 717

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 717

cgtacgtaaa gcggccgcgc tcatcatgga ctttccg

37

<210> 718

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 718

cgtacgtaaa gcggccgcgc actgctgttt gatgagg

37

<210> 719

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 719

cgtacgtaaa gcggccgcga gtgtggaaga aaggctgc

38

<210> 720

<211> 38

<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 720
cgtacgtaaa gcggccgcag tagagcttcc cctcctgc 38

<210> 721
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 721
cgtacgtaaa gcggccgcgt ccatcacgaa gtccaggt 38

<210> 722
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 722
cagcggcctg gaggatgc 18

<210> 723
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 723
cgggccagta gaggtttcg 19

<210> 724
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 724
catcagccgc gccttcatg 19

<210> 725

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 725
cctgcatgtt ggtgaagtac 20

<210> 726
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 726
gcccgaacc tctactggac cgac 24

<210> 727
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 727
gcccaccca tcacagttca catt 24

<210> 728
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 728
cagcggcctg gaggatgc 18

<210> 729
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 729
cggtacatg tactggacag ctgattagc 29

<210> 730
<211> 21
<212> DNA
<213> Homo sapiens

<400> 730
gtcagcctgg aggagttctc a 21

<210> 731
<211> 22
<212> DNA
<213> Homo sapiens

<400> 731
tcacccttgg caatacagat gt 22

<210> 732
<211> 21
<212> DNA
<213> Homo sapiens

<400> 732
cccacccatg tgcccgtgac a 21

<210> 733
<211> 19
<212> DNA
<213> Homo sapiens

<400> 733
cgtgattgcc gacgatctc 19

<210> 734
<211> 19
<212> DNA
<213> Homo sapiens

<400> 734
ttccggccgc tagtcttgt 19

<210> 735
<211> 26
<212> DNA
<213> Homo sapiens

<400> 735
cgcacccgtt cggctctgacg cagtac 26

<210> 736
<211> 22
<212> DNA
<213> Mouse

<400> 736

ctttcccccac gagtatgttg gt 22

<210> 737
<211> 19
<212> DNA
<213> Mouse

<400> 737
aagggaccgt gctgtgagc 19

<210> 738
<211> 29
<212> DNA
<213> Mouse

<400> 738
agcccctcat gtgcctctca acttcatag 29

<210> 739
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 739
gagcgggcag ggatggatgg 20

<210> 740
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 740
aggttggcac ggtggatgaa gc 22

<210> 741
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 741
aagcttggtt aaactgggca tggtggcaca tggttgtaat 40

<210> 742
<211> 34
<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 742

gggcttccac ccagtcagtc cagtacatgt acct

34

<210> 743

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 743

ctgactgggt ggaagcacc cggatcgagc

30

<210> 744

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 744

gaattcatcg gtacctgtgc ggccgcttca ttg

33

<210> 745

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 745

tgagatgtcc tgtctgtggc

20

<210> 746

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 746

tccttccttc cctacagttg

20

<210> 747

<211> 25

<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 747
cctaaggatc tccttgtgtc tgtgg 25

<210> 748
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 748
ctgcagcagg tcagtagcct gc 22

<210> 749
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 749
tgagatgtcc tgtctgtggc 20

<210> 750
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 750
tccttccttc cctacagttg 20

<210> 751
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 751
gcttctggcg tgtgaccggc g 21

<210> 752

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 752
gccgcacagc gccagcagca gc 22

<210> 753
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 753
cacccacgcc ccacagccag ta 22

<210> 754
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 754
atttgcctc ccatatgtcc ttcc 24

<210> 755
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 755
ttcctcccag ccctcctcca tcag 24

<210> 756
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 756
gccgcacagc gccagcagca gc 22

<210> 757
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 757
 gaatggcgcc cccgacgac 19

<210> 758
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 758
 gctcccatc atcagttcca tagg 24

<210> 759
 <211> 5278
 <212> DNA
 <213> Mouse

<400> 759
 tctagactcg aggcggccgc ccattgtgca ctaaagcgcc gccgccgcgc catggagccc 60
 gagtgcgcgc ggcgcggggc cgtccggccg ccggacaaca tggaggcagc gccgcccggg 120
 ccgccgtggc cgtctgtgct gctgctgctg ctgctgctgg cgtctgtgcg ctgccccggc 180
 cccgccgcgc cctgcgcgct cctgctatct gccaacgcgc gggacgtacg gctggtggac 240
 gccggcggag tcaagctgga gtccaccatc gtggtcagcg gcctggagga tgcggccgca 300
 gtggacttcc agttttccaa gggagccgtg tactggacag acgtgagcga ggaggccatc 360
 aagcagacct acctgaacca gacggggggc gccgtgcaga acgtggtcat ctccggcctg 420
 gtctctcccc acggcctcgc ctgcgactgg gtgggcaaga agctgtactg gacggactca 480
 gagaccaacc gcacgcgggt ggccaacctc aatggcacat cccggaagggt gctcttctgg 540
 caggaccttg accagcggag ggccatcgcc ttggaccccg ctcacgggta catgtactgg 600
 acagactggg ttgagacgcc ccggattgag cgggcaggga tggatggcag caccggaag 660
 atcattgtgg actcggacat ttactggccc aatggactga ccatcgacct ggaggagcag 720
 aagctctact gggctgacgc caagctcagc ttcatccacc gtgccaacct ggacggctcg 780
 ttccggcaga aggtgggtgga gggcagcctg acgcacccct tcgccctgac gctctccggg 840
 gacactctgt actggacaga ctggcagacc cgtccatcc atgcctgcaa caagcgcact 900
 ggggggaaga ggaaggagat cctgagtgcc ctctactcac ccatggacat ccagggtgctg 960
 agccaggagc ggcagccttt cttccacact cgtctgtagg aggacaatgg cggctgctcc 1020
 cacctgtgcc tgcgtgtccc aagcgagcct ttctacacat gcgcctgcc caccgggtgtg 1080
 cagctgcagg acaacggcag gacgtgtaag gcaggagccg aggaggtgct gctgctggcc 1140
 cggcggacgg acctacggag gatctcgtg gacacgccg acttcaccga catcgtgctg 1200
 cagggtggacg acatccggca cgccattgcc atcgactacg acccgctaga gggctatgtc 1260
 tactggacag atgacgaggt gcggggccatc cgcaggcggt acctggacgg gtctggggcg 1320
 cagacgctgg tcaacaccga gatcaacgac ccgatggca tcgcggtcga ctgggtggcc 1380
 cgaaacctct actggaccga cacgggcaag gaccgcacg aggtgacgcg cctcaacggc 1440
 acctcccgca agatcctggt gtcggaggac ctggacgagc cccgagccat cgcactgcac 1500
 cccgtgatgg gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 1560

aacttgatg	ggcaggagcg	gcgtgtgctg	gtcaatgcct	ccctcgggtg	gcccacggc	1620
ctggccctgg	acctgcagga	ggggaagctc	tactggggag	acgccaagac	agacaagatc	1680
gaggtgatca	atgttgatgg	gacgaagagg	cggaccctcc	tggaggacaa	gctcccgcac	1740
atcttcgggt	tcacgctgct	gggggacttc	atctactgga	ctgactggca	gcgcccgcagc	1800
atcgagcggg	tgcacaaggt	caaggccagc	cgggacgtca	tcattgacca	gctgcccgcac	1860
ctgatggggc	tcaaagctgt	gaatgtggcc	aaggtcgtcg	gaaccaaccc	gtgtgcggac	1920
aggaacgggg	ggtgcagcca	cctgtgcttc	ttcacacccc	acgcaacccg	gtgtggctgc	1980
cccacgcggc	tggagctgct	gagtgcacatg	aagacctgca	tcgtgcctga	ggccttcttg	2040
gtcttcacca	gcagagccgc	catccacagg	atctccctcg	agaccaataa	caacgcacgtg	2100
gccatcccgc	tcacgggcgt	caaggaggcc	tcagccctgg	actttgatgt	gtccaacaac	2160
cacatctact	ggacagacgt	cagcctgaag	accatcagcc	gcgccttcat	gaacgggagc	2220
tcggtggagc	acgtgggtga	gtttggcctt	gactaccccg	agggcatggc	cgttgactgg	2280
atgggcaaga	acctctactg	ggccgacact	gggaccaaca	gaatcgaagt	ggcgcggtcg	2340
gacgggcagt	tccggcaagt	cctcgtgtgg	agggacttgg	acaacccgag	gtcgtggcc	2400
ctggatccca	ccaagggcta	catctactgg	accgagtggg	gcggcaagcc	gaggatcgtg	2460
cgggccttca	tggacgggac	caactgcatg	acgctggtgg	acaagggtggg	ccgggccaac	2520
gacctcacca	ttgactacgc	tgaccagcgc	ctctactgga	ccgacctgga	caccaacatg	2580
atcgatcgct	ccaacatgct	gggtcaggag	cgggtcgtga	ttgccgacga	tctcccgcac	2640
ccgttcgggtc	tcagcagta	cagcgattat	atctactgga	cagactggaa	tctgcacagc	2700
attgagcggg	ccgacaagac	tagcgccggg	aaccgcaccc	tcacccaggg	ccacctggac	2760
ttcgtgatgg	acatcctggt	gttccactcc	tcccgcagg	atggcctcaa	tgactgtatg	2820
cacaacaacg	ggcagtgtgg	gcagctgtgc	cttgccatcc	ccggcgccca	ccgctgcggc	2880
tgccctcac	actacaccct	ggaccccagc	agccgcaact	gcagcccgc	caccaccttc	2940
ttgctgttca	gccagaaatc	tgccatcagt	cggatgatcc	cggacgacca	gcacagcccg	3000
gatctcatcc	tggccctgca	tggactgagg	aacgtcaaag	ccatcgacta	tgacctactg	3060
gacaagttca	tctactgggt	ggatgggcgc	cagaacatca	agcgagccaa	ggacgcaggg	3120
accagccct	ttgttttgac	ctctctgagc	caaggccaaa	accagacag	gcagcccac	3180
gacctcagca	tcgacatcta	cagccggaca	ctgttctgga	cgtgcgaggg	caccaatacc	3240
atcaacgtcc	acaggctgag	cggggaagcc	atgggggtgg	tgctgcgtgg	ggaccgcgac	3300
aagcccaggg	ccatcgctcg	caacgcggag	cgagggtacc	tgtacttcac	caacatgcag	3360
gaccgggcag	ccaagatcga	acgcgcagcc	ctggacggca	ccgagcgcga	ggctcctctc	3420
accaccggcc	tcacccgccc	tgtggccctg	gtggtagaca	acacactggg	caagctgttc	3480
tgggtggacg	cggacctgaa	gcgcattgag	agctgtgacc	tgtcaggggc	caaccgcctg	3540
accctggagg	acgccaacat	cgtgcagcct	ctgggcctga	ccatccttgg	caagcatctc	3600
tactggatcg	accgccagca	gcagatgatc	gagcgtgtgg	agaagaccac	cggggacaag	3660
cggactcgca	tccagggccg	tgctgcccac	ctcactggca	tccatgcagt	ggaggaagtc	3720
agcctggagg	agtctcagc	ccacctatgt	gcccgtagca	atggtggctg	ctcccacatc	3780
tgtattgcca	agggtgatgg	gacaccacgg	tgctcatgcc	cagtccacct	cgtgctcctg	3840
cagaacctgc	tgacctgtgg	agagccgccc	acctgctccc	cggaccagtt	tgcattgtgc	3900
acaggggaga	tcgactgtat	ccccggggcc	tggcgctgtg	acggctttcc	cgagtgcgat	3960
gaccagagcg	acgaggaggg	ctgccccgtg	tgctccgccc	cccagttccc	ctgcgcgagg	4020
ggtcagtgtg	tggacctgcg	cctgcgctgc	gacggcgagg	cagactgtca	ggaccgctca	4080
gacgaggcgg	actgtgacgc	catctgcctg	cccaaccagt	tccggtgtgc	gagcggccag	4140
tgtgtcctca	tcaaacagca	gtgcgactcc	ttccccgact	gtatcgacgg	ctccgacgag	4200
ctcatgtgtg	aatcaccaa	gccgcctca	gacgacagcc	cggcccacag	cagtgccatc	4260
gggcccgtca	ttggcatcat	cctctctctc	ttcgtcatgg	gtggtgtcta	ttttgtgtgc	4320
cagcgcgtgg	tgtgccagcg	ctatgcgggg	gccaacgggc	ccttcccgcg	cgagtatgtc	4380
agcgggaccc	cgcacgtgcc	cctcaatttc	atagccccgg	gcggttccca	gcattggccc	4440
ttcacaggca	tcgcatgcgg	aaagtccatg	atgagctccg	tgagcctgat	ggggggccgg	4500
ggcgggggtg	ccctctacga	ccggaaccac	gtcacagggg	cctcgtccag	cagctcgtcc	4560
agcacgaagg	ccacgctgta	cccgccgatc	ctgaacccgc	cgccctcccc	ggccacggac	4620
ccctccctgt	acaacatgga	catgttctac	tcttcaaaca	ttccggccac	tgcgagaccg	4680
tacaggccct	acatcatteg	aggaatggcg	cccccgacga	cgccctgcag	caccgcagtg	4740
tgtgacagcg	actacagcgc	cagccgctgg	aaggccagca	agtactacct	ggatttgaac	4800

```

tcggactcag acccctatcc acccccaccc acgccccaca gccagtacct gtcggcggag 4860
gacagctgcc cgccctcgcc cgccaccgag aggagctact tccatctctt cccgccccct 4920
ccgtccccct gcacggactc atcctgacct cggcggggcc actctggctt ctctgtgccc 4980
ctgtaaatag ttttaaatat gaacaaagaa aaaaatata tttatgattt aaaaaataaa 5040
tataattggg attttaaaaa catgagaaat gtgaactgtg atgggggtggg cagggctggg 5100
agaactttgt acagtggaga aatattttata aacttaattt tgtaaaacag aaaaaaaaaa 5160
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5220
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcggccgc 5278

```

<210> 760

<211> 5278

<212> DNA

<213> Homo sapiens

<400> 760

```

tctagactcg aggcgggcgc ccattgtgca ctaaagcgcc gccgccgcgc catggagccc 60
gagtgaagcg ggcgcggggc cgtccggcgc ccggacaaca tggaggcagc gccgcccggg 120
ccgccgtggc cgctgctgct gctgctgctg ctgctgctgg cgctgtgcgc ctgcccggcc 180
cccgccgcgc cctcgccgct cctgctattt gccaacgcgc gggacgtacg gctgggtggc 240
gccggcggag tcaagctgga gtccaccatc gtggtcagcg gcctggagga tgcggccgca 300
gtggacttcc agttttccaa gggagccgtg tactggacag acgtgagcga ggaggccatc 360
aagcagacct acctgaacca gacggggggc gccgtgcaga acgtggtcat ctccggcctg 420
gtctctcccg acggcctcgc ctgcgactgg gtgggcaaga agctgtactg gacggactca 480
gagaccaacc gcctcgaggt ggccaacctc aatggcacat cccggaaggt gctcttctgg 540
caggaccttg accagccgag ggccatcgcc ttggaccccg ctacgggta catgtactgg 600
acagactggg gtgagacgcc ccggattgag cgggcaggga tggatggcag caccggaag 660
atcattgtgg actcggacat ttactggccc aatggactga ccatcgacct ggaggagcag 720
aagctctact gggctgacgc caagctcagc ttcattccacc gtgccaacct ggacggctcg 780
ttccggcaga aggtgggtgga gggcagcctg acgcaccctc tcgccctgac gctctccggg 840
gacactctgt actggacaga ctggcagacc cgctccatcc atgcctgcaa caagcgcact 900
ggggggaaga ggaaggagat cctgagtgcc ctctactcac ccatggacat ccaggtgctg 960
agccaggagc ggcagccttt cttccacact cgctgtgagg aggacaatgg cggctgctcc 1020
cacctgtgcc tgctgtcccc aagcgagcct ttctacacat gcgcctgccc cacgggtgtg 1080
cagctgcagg acaacggcag gacgtgtaag gcaggagccg aggaggtgct gctgctggcc 1140
cggcggacgg acctacggag gatctcgctg gacacgcccg acctcaccga catcgtgctg 1200
caggtggacg acatccggca cgccattgcc atcgactacg acccgctaga gggctatgtc 1260
tactggacag atgacgaggt gcggggccat cgcaggcgct acctggacgg gtctggggcg 1320
cagacgctgt tcaacacgga gatcaacgac cccgatggca tcgcggtcga ctgggtggcc 1380
cgaaacctct actggaccga cacgggcacg gaccgcatcg aggtgacgcg cctcaacggc 1440
acctcccgca agatcctggt gtcggaggac ctggacgagc cccgagccat cgactgcac 1500
cccgtgatgg gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 1560
aacttggatg ggcaggagcg gcgtgtgctg gtcaatgcct ccctcgggtg gcccacggc 1620
ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 1680
gaggtgatca atgttgatgg gacgaagagg cggaccctcc tggaggacaa gctcccgcac 1740
attttcgggt tcacgctgct gggggacttc atctactgga ctgactggca gcgcgcgacg 1800
atcgagcggg tgcacaaggt caaggccagc cgggacgtca tcattgacca gctgcccgcac 1860
ctgatggggc tcaaagctgt gaatgtggcc aaggtcgtcg gaaccaacct gtgtgcggac 1920
aggaacgggg ggtgcagcca cctgtgcttc ttcacacccc acgcaacctg gtgtggctgc 1980
cccatcggcc tggagctgct gagtgcacatg aagacctgca tcgtgcctga ggccttcttg 2040
gtcttcacca gcagagccgc catccacagg atctccctcg agaccaataa caacgacgtg 2100
gccatccgcg tcacgggcgt caaggaggcc tcagccctgg actttgatgt gtccaacaac 2160
cacatctact ggacagacgt cagcctgaag accatcagcc gcgccttcat gaacgggagc 2220
tcggtggagc acgtgggtgga gtttggcctt gactaccccg agggcatggc cgttgactgg 2280
atgggcaaga acctctactg ggccgacact gggaccaaca gaatcgaagt ggcgcggctg 2340

```

```

gacgggcagt tccggcaagt cctcgtgtgg agggacttgg acaaccgag gtcgctggcc 2400
ctggatccca ccaagggcta catctactgg accgagtggg gcggaagcc gaggatcgtg 2460
cgggccttca tggacgggac caactgcatg acgctggtgg acaagggtggg cggggccaac 2520
gacctacca ttgactacgc tgaccagcgc ctctactgga cggacctgga caccaacatg 2580
atcgagtcgt ccaacatgct gggtcaggag cgggtcgtga ttgccgacga tctcccgac 2640
ccgttcggtc tgacgcagta cagcgattat atctactgga cagactggaa tctgcacagc 2700
attgagcggg ccgacaagac tagcggccgg aaccgcaccc tcatccaggg ccacctggac 2760
ttcgtgatgg acatcctggg gttccactcc tcccgccagg atggcctcaa tgactgtatg 2820
cacaacaacg ggcagtgtgg gcagctgtgc cttgccatcc ccggcgcca ccgctgcggc 2880
tgcgccctac actacacctt ggaccccagc agccgcaact gcagcccgc caccaccttc 2940
ttgctgttca gccagaaatc tgccatcagt cggatgatcc cggacgacca gcacagcccg 3000
gatctcatcc tgcccctgca tggactgagg aacgtcaaa ccatcgacta tgacctactg 3060
gacaagttca tctactgggt ggatggggcg cagaacatca agcgagccaa ggacgacggg 3120
accagccct ttgttttgac ctctctgagc caaggccaaa accagacag gcagcccac 3180
gacctagca tcgacatcta cagccggaca ctgttctgga cgtgcgaggc caccaatacc 3240
atcaacgtcc acaggctgag cggggaagcc atgggggtgg tgcgtcgtgg ggaccgcgac 3300
aagcccaggg ccatcgctgt caacgcggag cgagggtacc tgtacttcac caacatgcag 3360
gaccgggag ccaagatcga acgcgagcc ctggacggca ccgagcgoga ggtcctcttc 3420
accacgggcc tcatccggcc tgtggccctg gtggtagaca acacactggg caagctgttc 3480
tgggtggagc cggacctgaa gcgcattgag agctgtgacc tgtcaggggg caaccgcctc 3540
accctggagg accccaacat cgtgcagcct ctgggcctga ccatccttgg caagcatctc 3600
tactggatcg accgccagca gcagatgata gagcgtgtgg agaagaccac cgggggacaag 3660
cggactcgca tccagggccg tgcgcccac ctactggca tccatgcagt ggaggaaagt 3720
agcctggagg agttctcagc ccacccatgt gcccgtaga atggtggctg ctcccacatc 3780
tgtattgcca aggggtgatg gacaccacgg tgcctatgcc cagtcacact cgtgctcctg 3840
cagaacctgc tgacctgtgg agagccgccc acctgctccc cggaccagtt tgcattgtgc 3900
acaggggaga tcgactgtat ccccggggcc tggcgtgtg acggctttcc cgagtgcgat 3960
gaccagagcg acgaggagg ctgcccctgt tgcctcgccg ccagttccc ctgcgcgagg 4020
ggtcagtgtg tggacctgcg cctgcgctgc gacggcgagg cagactgtca ggaccgctca 4080
gacgaggcgg actgtgacgc catctgcctg cccaaccagt tccggtgtgc gagcgccag 4140
tgtgtcctca tcaaacagca gtgcgactcc tccccgact gtatcgacgg ctccgacgag 4200
ctcatgtgtg aaatcaccaa gccgcctca gacgacagcc cggccacag cagtgccatc 4260
gggcccgtca ttggcatcat cctctctctc ttcgctcatg gtggtgtcta ttttgtgtgc 4320
cagcgcgtgg tgtgccagcg ctatgcgggg gccaacgggc ccttcccga cgagtatgtc 4380
agcgggaccc cgcacgtgcc cctcaatttc atagcccgg gcggttcca gcattggccc 4440
ttcacaggca tcgcatgcgg aaagtccatg atgagctcc tgagcctgat ggggggcccg 4500
ggcggggtgc ccctctacga ccggaaccac gtcacagggg cctcgtccag cagctcgtcc 4560
agcacgaagg ccacgctgta cccgccgatc ctgaaccgcg cgcctcccc ggccacggac 4620
ccctccctgt acaacatgga catgttctac tcttcaaaca ttccggccac tgcgagaccg 4680
tacaggccct acatcattcg aggaatggcg ccccgacga agcctgcag caccgacgtg 4740
tgtgacagcg actacagcgc cagccgctgg aaggccagca agtactacct ggatttgaac 4800
tcggactcag acccctatcc acccccaccc acgccccaca gccagtacct gtcggcgagg 4860
gacagctgcc cgcctcgcg cgcaccgag agggagctact tccatctctt cccgccccct 4920
ccgtccccct gcacggactc atcctgacct cggccggggc actctggctt ctctgtgccc 4980
ctgtaaatag ttttaaata gaacaaagaa aaaaatata tttatgattt aaaaaataaa 5040
tataattggg attttaaaa catgagaaat gtgaactgtg atggggtggg cagggtctggg 5100
agaactttgt acagtggaga aatatattata aacttaattt tgtaaaacag aaaaaaaaaa 5160
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5220
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcggccgc 5278

```

<210> 761

<211> 1615

<212> PRT

<213> Homo sapiens

<400> 761

```

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1          5          10          15
Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
          20          25          30
Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
          35          40          45
Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
          50          55          60
Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
65          70          75          80
Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
          85          90          95
Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
          100          105          110
Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
          115          120          125
Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
          130          135          140
Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala Leu Asp Pro
145          150          155          160
Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
          165          170          175
Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
          180          185          190
Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
          195          200          205
Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
          210          215          220
Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
225          230          235          240
Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
          245          250          255
Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
          260          265          270
Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
          275          280          285
Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
          290          295          300
Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
305          310          315          320
Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly Arg Thr Cys
          325          330          335
Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu
          340          345          350
Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
          355          360          365
Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
          370          375          380
Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
385          390          395          400
Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
          405          410          415
Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp

```

-285-

850	855	860
Leu His Ser Ile Glu Arg	Ala Asp Lys Thr Ser	Gly Arg Asn Arg Thr
865	870	875
Leu Ile Gln Gly His Leu	Asp Phe Val Met Asp	Ile Leu Val Phe His
885	890	895
Ser Ser Arg Gln Asp Gly	Leu Asn Asp Cys Met	His Asn Asn Gly Gln
900	905	910
Cys Gly Gln Leu Cys Leu	Ala Ile Pro Gly Gly	His Arg Cys Gly Cys
915	920	925
Ala Ser His Tyr Thr Leu	Asp Pro Ser Ser Arg	Asn Cys Ser Pro Pro
930	935	940
Thr Thr Phe Leu Leu Phe	Ser Gln Lys Ser Ala	Ile Ser Arg Met Ile
945	950	955
Pro Asp Asp Gln His Ser	Pro Asp Leu Ile Leu	Pro Leu His Gly Leu
965	970	975
Arg Asn Val Lys Ala Ile	Asp Tyr Asp Pro Leu	Asp Lys Phe Ile Tyr
980	985	990
Trp Val Asp Gly Arg Gln	Asn Ile Lys Arg Ala	Lys Asp Asp Gly Thr
995	1000	1005
Gln Pro Phe Val Leu Thr	Ser Leu Ser Gln Gly	Gln Asn Pro Asp Arg
1010	1015	1020
Gln Pro His Asp Leu Ser	Ile Asp Ile Tyr Ser	Arg Thr Leu Phe Trp
1025	1030	1035
Thr Cys Glu Ala Thr Asn	Thr Ile Asn Val His	Arg Leu Ser Gly Glu
1045	1050	1055
Ala Met Gly Val Val Leu	Arg Gly Asp Arg Asp	Lys Pro Arg Ala Ile
1060	1065	1070
Val Val Asn Ala Glu Arg	Gly Tyr Leu Tyr Phe	Thr Asn Met Gln Asp
1075	1080	1085
Arg Ala Ala Lys Ile Glu	Arg Ala Ala Leu Asp	Gly Thr Glu Arg Glu
1090	1095	1100
Val Leu Phe Thr Thr Gly	Leu Ile Arg Pro Val	Ala Leu Val Val Asp
1105	1110	1115
Asn Thr Leu Gly Lys Leu	Phe Trp Val Asp Ala	Asp Leu Lys Arg Ile
1125	1130	1135
Glu Ser Cys Asp Leu Ser	Gly Ala Asn Arg Leu	Thr Leu Glu Asp Ala
1140	1145	1150
Asn Ile Val Gln Pro Leu	Gly Leu Thr Ile Leu	Gly Lys His Leu Tyr
1155	1160	1165
Trp Ile Asp Arg Gln Gln	Gln Met Ile Glu Arg	Val Glu Lys Thr Thr
1170	1175	1180
Gly Asp Lys Arg Thr Arg	Ile Gln Gly Arg Val	Ala His Leu Thr Gly
1185	1190	1195
Ile His Ala Val Glu Glu	Val Ser Leu Glu Glu	Phe Ser Ala His Pro
1205	1210	1215
Cys Ala Arg Asp Asn Gly	Gly Cys Ser His Ile	Cys Ile Ala Lys Gly
1220	1225	1230
Asp Gly Thr Pro Arg Cys	Ser Cys Pro Val His	Leu Val Leu Leu Gln
1235	1240	1245
Asn Leu Leu Thr Cys Gly	Glu Pro Pro Thr Cys	Ser Pro Asp Gln Phe
1250	1255	1260
Ala Cys Ala Thr Gly Glu	Ile Asp Cys Ile Pro	Gly Ala Trp Arg Cys
1265	1270	1275
Asp Gly Phe Pro Glu Cys	Asp Asp Gln Ser Asp	Glu Glu Gly Cys Pro

```
<210> 762
<211> 1564
<212> PRT
<213> Mouse
```

<400> 762

Met	Glu	Thr	Ala	Pro	Thr	Arg	Ala	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Leu
1				5				10					15		
Leu	Leu	Leu	Val	Leu	Tyr	Cys	Ser	Leu	Val	Pro	Ala	Ala	Ala	Ser	Pro
			20					25					30		
Leu	Leu	Leu	Phe	Ala	Asn	Arg	Arg	Asp	Val	Arg	Leu	Val	Asp	Ala	Gly

-288-

465					470					475				480
Pro	Lys	Ile	Glu	Cys	Ala	Asn	Leu	Asp	Gly	Arg	Asp	Arg	His	Val
				485					490					495
Val	Asn	Thr	Ser	Leu	Gly	Trp	Pro	Asn	Gly	Leu	Ala	Leu	Asp	Leu
			500					505					510	
Glu	Gly	Lys	Leu	Tyr	Trp	Gly	Asp	Ala	Lys	Thr	Asp	Lys	Ile	Glu
		515				520					525			Val
Ile	Asn	Ile	Asp	Gly	Thr	Lys	Arg	Lys	Thr	Leu	Leu	Glu	Asp	Lys
	530					535				540				Leu
Pro	His	Ile	Phe	Gly	Phe	Thr	Leu	Leu	Gly	Asp	Phe	Ile	Tyr	Trp
545					550					555				560
Asp	Trp	Gln	Arg	Arg	Ser	Ile	Glu	Arg	Val	His	Lys	Val	Lys	Ala
			565						570					575
Arg	Asp	Val	Ile	Asp	Gln	Leu	Pro	Asp	Leu	Met	Gly	Leu	Lys	Ala
		580					585					590		
Val	Asn	Val	Ala	Lys	Val	Val	Gly	Thr	Asn	Pro	Cys	Ala	Asp	Gly
		595					600					605		Asn
Gly	Gly	Cys	Ser	His	Leu	Cys	Phe	Phe	Thr	Pro	Arg	Ala	Thr	Lys
	610					615					620			Cys
Gly	Cys	Pro	Ile	Gly	Leu	Glu	Leu	Leu	Ser	Asp	Met	Lys	Thr	Cys
625					630					635				Ile
Ile	Pro	Glu	Ala	Phe	Leu	Val	Phe	Thr	Ser	Arg	Ala	Thr	Ile	His
			645						650					655
Ile	Ser	Leu	Glu	Thr	Asn	Asn	Asn	Asp	Val	Ala	Ile	Pro	Leu	Thr
		660						665					670	Gly
Val	Lys	Glu	Ala	Ser	Ala	Leu	Asp	Phe	Asp	Val	Ser	Asn	Asn	His
		675					680					685		Ile
Tyr	Trp	Thr	Asp	Val	Ser	Leu	Lys	Thr	Ile	Ser	Arg	Ala	Phe	Met
	690					695					700			Asn
Gly	Ser	Ser	Val	Glu	His	Val	Ile	Glu	Phe	Gly	Leu	Asp	Tyr	Pro
705					710					715				Glu
Gly	Met	Ala	Val	Asp	Trp	Met	Gly	Lys	Asn	Leu	Tyr	Trp	Ala	Asp
			725						730					Thr
Gly	Thr	Asn	Arg	Ile	Glu	Val	Ala	Arg	Leu	Asp	Gly	Gln	Phe	Arg
		740						745				750		Gln
Val	Leu	Val	Trp	Arg	Asp	Leu	Asp	Asn	Pro	Arg	Ser	Leu	Ala	Leu
		755					760					765		Asp
Pro	Thr	Lys	Gly	Tyr	Ile	Tyr	Trp	Thr	Glu	Trp	Gly	Gly	Lys	Pro
	770					775					780			Arg
Ile	Val	Arg	Ala	Phe	Met	Asp	Gly	Thr	Asn	Cys	Met	Thr	Leu	Val
785					790					795				Asp
Lys	Val	Gly	Arg	Ala	Asn	Asp	Leu	Thr	Ile	Asp	Tyr	Ala	Asp	Gln
			805						810					Arg
Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Met	Ile	Glu	Ser	Ser	Asn
		820					825					830		Met
Leu	Gly	Gln	Glu	Arg	Met	Val	Ile	Ala	Asp	Asp	Leu	Pro	Tyr	Pro
		835					840				845			Phe
Gly	Leu	Thr	Gln	Tyr	Ser	Asp	Tyr	Ile	Tyr	Trp	Thr	Asp	Trp	Asn
	850					855					860			Leu
His	Ser	Ile	Glu	Arg	Ala	Asp	Lys	Thr	Ser	Gly	Arg	Asn	Arg	Thr
865					870					875				Leu
Ile	Gln	Gly	His	Leu	Asp	Phe	Val	Met	Asp	Ile	Leu	Val	Phe	His
			885						890					Ser
Ser	Arg	Gln	Asp	Gly	Leu	Asn	Asp	Cys	Val	His	Ser	Asn	Gly	Gln
														Cys

900					905					910					
Gly	Gln	Leu	Cys	Leu	Ala	Ile	Pro	Gly	Gly	His	Arg	Cys	Gly	Cys	Ala
915					920					925					
Ser	His	Tyr	Thr	Leu	Asp	Pro	Ser	Ser	Arg	Asn	Cys	Ser	Pro	Pro	Ser
930					935					940					
Thr	Phe	Leu	Leu	Phe	Ser	Gln	Lys	Phe	Ala	Ile	Ser	Arg	Met	Ile	Pro
945					950					955					
Asp	Asp	Gln	Leu	Ser	Pro	Asp	Leu	Val	Leu	Pro	Leu	His	Gly	Leu	Arg
965					970					975					
Asn	Val	Lys	Ala	Ile	Asn	Tyr	Asp	Pro	Leu	Asp	Lys	Phe	Ile	Tyr	Trp
980					985					990					
Val	Asp	Gly	Arg	Gln	Asn	Ile	Lys	Arg	Ala	Lys	Asp	Asp	Gly	Thr	Gln
995					1000					1005					
Pro	Ser	Met	Leu	Thr	Ser	Pro	Ser	Gln	Ser	Leu	Ser	Pro	Asp	Arg	Gln
1010					1015					1020					
Pro	His	Asp	Leu	Ser	Ile	Asp	Ile	Tyr	Ser	Arg	Thr	Leu	Phe	Trp	Thr
1025					1030					1035					
Cys	Glu	Ala	Thr	Asn	Thr	Ile	Asn	Val	His	Arg	Leu	Asp	Gly	Asp	Ala
1045					1050					1055					
Met	Gly	Val	Val	Leu	Arg	Gly	Asp	Arg	Asp	Lys	Pro	Arg	Ala	Ile	Ala
1060					1065					1070					
Val	Asn	Ala	Glu	Arg	Gly	Tyr	Met	Tyr	Phe	Thr	Asn	Met	Gln	Asp	His
1075					1080					1085					
Ala	Ala	Lys	Ile	Glu	Arg	Ala	Ser	Leu	Asp	Gly	Thr	Glu	Arg	Glu	Val
1090					1095					1100					
Leu	Phe	Thr	Thr	Gly	Leu	Ile	Arg	Pro	Val	Ala	Leu	Val	Val	Asp	Asn
1105					1110					1115					
Ala	Leu	Gly	Lys	Leu	Phe	Trp	Val	Asp	Ala	Asp	Leu	Lys	Arg	Ile	Glu
1125					1130					1135					
Ser	Cys	Asp	Leu	Ser	Gly	Ala	Asn	Arg	Leu	Thr	Leu	Glu	Asp	Ala	Asn
1140					1145					1150					
Ile	Val	Gln	Pro	Val	Gly	Leu	Thr	Val	Leu	Gly	Arg	His	Leu	Tyr	Trp
1155					1160					1165					
Ile	Asp	Arg	Gln	Gln	Gln	Met	Ile	Glu	Arg	Val	Glu	Lys	Thr	Thr	Gly
1170					1175					1180					
Asp	Lys	Arg	Thr	Arg	Val	Gln	Gly	Arg	Val	Thr	His	Leu	Thr	Gly	Ile
1185					1190					1195					
His	Ala	Val	Glu	Glu	Val	Ser	Leu	Glu	Glu	Phe	Ser	Ala	His	Pro	Cys
1205					1210					1215					
Ala	Arg	Asp	Asn	Gly	Gly	Cys	Ser	His	Ile	Cys	Ile	Ala	Lys	Gly	Asp
1220					1225					1230					
Gly	Thr	Pro	Arg	Cys	Ser	Cys	Pro	Val	His	Leu	Val	Leu	Leu	Gln	Asn
1235					1240					1245					
Leu	Leu	Thr	Cys	Gly	Glu	Pro	Pro	Thr	Cys	Ser	Pro	Asp	Gln	Phe	Ala
1250					1255					1260					
Cys	Thr	Thr	Gly	Glu	Ile	Asp	Cys	Ile	Pro	Gly	Ala	Trp	Arg	Cys	Asp
1265					1270					1275					
Gly	Phe	Pro	Glu	Cys	Ala	Asp	Gln	Ser	Asp	Glu	Glu	Gly	Cys	Pro	Val
1285					1290					1295					
Cys	Ser	Ala	Ser	Gln	Phe	Pro	Cys	Ala	Arg	Gly	Gln	Cys	Val	Asp	Leu
1300					1305					1310					
Arg	Leu	Arg	Cys	Asp	Gly	Glu	Ala	Asp	Cys	Gln	Asp	Arg	Ser	Asp	Glu
1315					1320					1325					
Ala	Asn	Cys	Asp	Ala	Val	Cys	Leu	Pro	Asn	Gln	Phe	Arg	Cys	Thr	Ser

1330 1335 1340
 Gly Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln Arg Val Met
 1345 1350 1355 1360
 Cys Gln Arg Tyr Thr Gly Ala Ser Gly Pro Phe Pro His Glu Tyr Val
 1365 1370 1375
 Gly Gly Ala Pro His Val Pro Leu Asn Phe Ile Ala Pro Gly Gly Ser
 1380 1385 1390
 Gln His Gly Pro Phe Pro Gly Ile Pro Cys Ser Lys Ser Val Met Ser
 1395 1400 1405
 Ser Met Ser Leu Val Gly Gly Arg Gly Ser Val Pro Leu Tyr Asp Arg
 1410 1415 1420
 Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser Ser Thr Lys Ala
 1425 1430 1435 1440
 Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro Ala Thr Asp
 1445 1450 1455
 Pro Ser Leu Tyr Asn Val Asp Val Phe Tyr Ser Ser Gly Ile Pro Ala
 1460 1465 1470
 Thr Ala Arg Pro Tyr Arg Pro Tyr Val Ile Arg Gly Met Ala Pro Pro
 1475 1480 1485
 Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr Ser Ile Ser
 1490 1495 1500
 Arg Trp Lys Ser Ser Lys Tyr Tyr Leu Asp Leu Asn Ser Asp Ser Asp
 1505 1510 1515 1520
 Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu Ser Ala Glu
 1525 1530 1535
 Asp Ser Cys Pro Pro Ser Pro Gly Thr Glu Arg Ser Tyr Cys His Leu
 1540 1545 1550
 Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
 1555 1560

<210> 763

<211> 1665

<212> PRT

<213> Homo sapiens

<400> 763

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
 20 25 30
 Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45
 Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60
 Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80
 Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95
 Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110
 Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125
 Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val

130		135		140
Leu Phe Trp Gln Asp	Leu Asp Gln Pro Arg Ala	Ile Ala Leu Asp Pro		
145	150	155	160	
Ala His Gly Tyr Met	Tyr Trp Thr Asp Trp Gly	Glu Thr Pro Arg Ile		
	165	170	175	
Glu Arg Ala Gly Met	Asp Gly Ser Thr Arg Lys	Ile Ile Val Asp Ser		
	180	185	190	
Asp Ile Tyr Trp Pro	Asn Gly Leu Thr Ile Asp	Leu Glu Glu Gln Lys		
	195	200	205	
Leu Tyr Trp Ala Asp	Ala Lys Leu Ser Phe Ile	His Arg Ala Asn Leu		
	210	215	220	
Asp Gly Ser Phe Arg	Gln Lys Val Val Glu Gly	Ser Leu Thr His Pro		
225	230	235	240	
Phe Ala Leu Thr Leu	Ser Gly Asp Thr Leu Tyr	Trp Thr Asp Trp Gln		
	245	250	255	
Thr Arg Ser Ile His	Ala Cys Asn Lys Arg Thr	Gly Gly Lys Arg Lys		
	260	265	270	
Glu Ile Leu Ser Ala	Leu Tyr Ser Pro Met Asp	Ile Gln Val Leu Ser		
	275	280	285	
Gln Glu Arg Gln Pro	Phe Phe His Thr Arg Cys	Glu Glu Asp Asn Gly		
	290	295	300	
Gly Cys Ser His Leu	Cys Leu Leu Ser Pro Ser	Glu Pro Phe Tyr Thr		
305	310	315	320	
Cys Ala Cys Pro Thr	Gly Val Gln Leu Gln Asp	Asn Gly Arg Thr Cys		
	325	330	335	
Lys Ala Gly Ala Glu	Glu Val Leu Leu Leu Ala	Arg Arg Thr Asp Leu		
	340	345	350	
Arg Arg Ile Ser Leu	Asp Thr Pro Asp Phe Thr	Asp Ile Val Leu Gln		
	355	360	365	
Val Asp Asp Ile Arg	His Ala Ile Ala Ile Asp	Tyr Asp Pro Leu Glu		
	370	375	380	
Gly Tyr Val Tyr Trp	Thr Asp Asp Glu Val Arg	Ala Ile Arg Arg Ala		
385	390	395	400	
Tyr Leu Asp Gly Ser	Gly Ala Gln Thr Leu Val	Asn Thr Glu Ile Asn		
	405	410	415	
Asp Pro Asp Gly Ile	Ala Val Asp Trp Val Ala	Arg Asn Leu Tyr Trp		
	420	425	430	
Thr Asp Thr Gly Thr	Asp Arg Ile Glu Val Thr	Arg Leu Asn Gly Thr		
	435	440	445	
Ser Arg Lys Ile Leu	Val Ser Glu Asp Leu Asp	Glu Pro Arg Ala Ile		
	450	455	460	
Ala Leu His Pro Val	Met Gly Leu Met Tyr Trp	Thr Asp Trp Gly Glu		
465	470	475	480	
Asn Pro Lys Ile Glu	Cys Ala Asn Leu Asp Gly	Gln Glu Arg Arg Val		
	485	490	495	
Leu Val Asn Ala Ser	Leu Gly Trp Pro Asn Gly	Leu Ala Leu Asp Leu		
	500	505	510	
Gln Glu Gly Lys Leu	Tyr Trp Gly Asp Ala Lys	Thr Asp Lys Ile Glu		
	515	520	525	
Val Ile Asn Val Asp	Gly Thr Lys Arg Arg Thr	Leu Leu Glu Asp Lys		
	530	535	540	
Leu Pro His Ile Phe	Gly Phe Thr Leu Leu Gly	Asp Phe Ile Tyr Trp		
545	550	555	560	
Thr Asp Trp Gln Arg	Arg Ser Ile Glu Arg Val	His Lys Val Lys Ala		

-293-

995	1000	1005
Met Val Ile Asp Glu Gln Gln Ser Pro Asp Ile Ile Leu Pro Ile His		
1010	1015	1020
Ser Leu Arg Asn Val Arg Ala Ile Asp Tyr Asp Pro Leu Asp Lys Gln		
1025	1030	1035
Leu Tyr Trp Ile Asp Ser Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp		1040
1045	1050	1055
Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro		
1060	1065	1070
Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu		
1075	1080	1085
Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser		
1090	1095	1100
Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg		
1105	1110	1115
Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met		1120
1125	1130	1135
Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu		
1140	1145	1150
Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val		
1155	1160	1165
Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys		
1170	1175	1180
Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu		
1185	1190	1195
Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His		1200
1205	1210	1215
Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys		
1220	1225	1230
Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu		
1235	1240	1245
Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala		
1250	1255	1260
His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala		
1265	1270	1275
Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu		1280
1285	1290	1295
Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp		
1300	1305	1310
Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp		
1315	1320	1325
Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly		
1330	1335	1340
Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys		
1345	1350	1355
Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg		1360
1365	1370	1375
Ser Asp Glu Ala Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg		
1380	1385	1390
Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe		
1395	1400	1405
Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys		
1410	1415	1420
Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val		

1425 1430 1435 1440
 Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val
 1445 1450 1455
 Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe
 1460 1465 1470
 Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile
 1475 1480 1485
 Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly
 1490 1495 1500
 Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val
 1505 1510 1515 1520
 Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser
 1525 1530 1535
 Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro
 1540 1545 1550
 Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser
 1555 1560 1565
 Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg
 1570 1575 1580
 Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser
 1585 1590 1595 1600
 Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu
 1605 1610 1615
 Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln
 1620 1625 1630
 Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg
 1635 1640 1645
 Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser
 1650 1655 1660
 Ser
 1665

<210> 764
 <211> 1613
 <212> PRT
 <213> Homo sapiens

<400> 764
 Met Gly Ala Val Leu Arg Ser Leu Leu Ala Cys Ser Phe Cys Val Leu
 1 5 10 15
 Leu Arg Ala Ala Pro Leu Leu Leu Tyr Ala Asn Arg Arg Asp Leu Arg
 20 25 30
 Leu Val Asp Ala Thr Asn Gly Lys Glu Asn Ala Thr Ile Val Val Gly
 35 40 45
 Gly Leu Glu Asp Ala Ala Ala Val Asp Phe Val Phe Ser His Gly Leu
 50 55 60
 Ile Tyr Trp Ser Asp Val Ser Glu Glu Ala Ile Lys Arg Thr Glu Phe
 65 70 75 80
 Asn Lys Thr Glu Ser Val Gln Asn Val Val Val Ser Gly Leu Leu Ser
 85 90 95
 Pro Asp Gly Leu Ala Cys Asp Trp Leu Gly Glu Lys Leu Tyr Trp Thr
 100 105 110
 Asp Ser Glu Thr Asn Arg Ile Glu Val Ser Asn Leu Asp Gly Ser Leu

115	120	125
Arg Lys Val Leu Phe Trp	Gln Glu Leu Asp Gln	Pro Arg Ala Ile Ala
130	135	140
Leu Asp Pro Ser Ser Gly	Phe Met Tyr Trp Thr	Asp Trp Gly Glu Val
145	150	155
Pro Lys Ile Glu Arg Ala	Gly Met Asp Gly Ser	Ser Arg Phe Ile Ile
165	170	175
Ile Asn Ser Glu Ile Tyr Trp	Pro Asn Gly Leu Thr	Leu Asp Tyr Glu
180	185	190
Glu Gln Lys Leu Tyr Trp	Ala Asp Ala Lys Leu	Asn Phe Ile His Lys
195	200	205
Ser Asn Leu Asp Gly Thr	Asn Arg Gln Ala Val	Val Lys Gly Ser Leu
210	215	220
Pro His Pro Phe Ala Leu	Thr Leu Phe Glu Asp	Ile Leu Tyr Trp Thr
225	230	235
Asp Trp Ser Thr His Ser	Ile Leu Ala Cys Asn	Lys Tyr Thr Gly Glu
245	250	255
Gly Leu Arg Glu Ile His	Ser Asp Ile Phe Ser	Pro Met Asp Ile His
260	265	270
Ala Phe Ser Gln Gln Arg	Gln Pro Asn Ala Thr	Asn Pro Cys Gly Ile
275	280	285
Asp Asn Gly Gly Cys Ser	His Leu Cys Leu Met	Ser Pro Val Lys Pro
290	295	300
Phe Tyr Gln Cys Ala Cys	Pro Thr Gly Val Lys	Leu Leu Glu Asn Gly
305	310	315
Lys Thr Cys Lys Asp Gly	Ala Thr Glu Leu Leu	Leu Leu Ala Arg Arg
325	330	335
Thr Asp Leu Arg Arg Ile	Ser Leu Asp Thr Pro	Asp Phe Thr Asp Ile
340	345	350
Val Leu Gln Leu Glu Asp	Ile Arg His Ala Ile	Ala Ile Asp Tyr Asp
355	360	365
Pro Val Glu Gly Tyr Ile	Tyr Trp Thr Asp Asp	Glu Val Arg Ala Ile
370	375	380
Arg Arg Ser Phe Ile Asp	Gly Ser Gly Ser Gln	Phe Val Val Thr Ala
385	390	395
Gln Ile Ala His Pro Asp	Gly Ile Ala Val Asp	Trp Val Ala Arg Asn
405	410	415
Leu Tyr Trp Thr Asp Thr	Gly Thr Asp Arg Ile	Glu Val Thr Arg Leu
420	425	430
Asn Gly Thr Met Arg Lys	Ile Leu Ile Ser Glu	Asp Leu Glu Glu Pro
435	440	445
Arg Ala Ile Val Leu Asp	Pro Met Val Gly Tyr	Met Tyr Trp Thr Asp
450	455	460
Trp Gly Glu Ile Pro Lys	Ile Glu Arg Ala Ala	Leu Asp Gly Ser Asp
465	470	475
Arg Val Val Leu Val Asn	Thr Ser Leu Gly Trp	Pro Asn Gly Leu Ala
485	490	495
Leu Asp Tyr Asp Glu Gly	Lys Ile Tyr Trp Gly	Asp Ala Lys Thr Asp
500	505	510
Lys Ile Glu Val Met Asn	Thr Asp Gly Thr Gly	Arg Arg Val Leu Val
515	520	525
Glu Asp Lys Ile Pro His	Ile Phe Gly Phe Thr	Leu Leu Gly Asp Tyr
530	535	540
Val Tyr Trp Thr Asp Trp	Gln Arg Arg Ser Ile	Glu Arg Val His Lys

545				550					555				560
Arg	Ser	Ala	Glu	Arg	Glu	Val	Ile	Ile	Asp	Gln	Leu	Pro	Asp
				565					570				575
Gly	Leu	Lys	Ala	Thr	Asn	Val	His	Arg	Val	Ile	Gly	Ser	Asn
				580					585				590
Ala	Glu	Glu	Asn	Gly	Gly	Cys	Ser	His	Leu	Cys	Leu	Tyr	Arg
				595					600				605
Gly	Leu	Arg	Cys	Ala	Cys	Pro	Ile	Gly	Phe	Glu	Leu	Ile	Ser
				610					615				620
Lys	Thr	Cys	Ile	Val	Pro	Glu	Ala	Phe	Leu	Leu	Phe	Ser	Arg
				625					630				635
Asp	Ile	Arg	Arg	Ile	Ser	Leu	Glu	Thr	Asn	Asn	Asn	Asn	Val
				645					650				655
Pro	Leu	Thr	Gly	Val	Lys	Glu	Ala	Ser	Ala	Leu	Asp	Phe	Asp
				660					665				670
Asp	Asn	Arg	Ile	Tyr	Trp	Thr	Asp	Ile	Ser	Leu	Lys	Thr	Ile
				675					680				685
Ala	Phe	Met	Asn	Gly	Ser	Ala	Leu	Glu	His	Val	Val	Glu	Phe
				690					695				700
Asp	Tyr	Pro	Glu	Gly	Met	Ala	Val	Asp	Trp	Leu	Gly	Lys	Asn
				705					710				715
Trp	Ala	Asp	Thr	Gly	Thr	Asn	Arg	Ile	Glu	Val	Ser	Lys	Leu
				725					730				735
Gln	His	Arg	Gln	Val	Leu	Val	Trp	Lys	Asp	Leu	Asp	Ser	Pro
				740					745				750
Leu	Ala	Leu	Asp	Pro	Ala	Glu	Gly	Phe	Met	Tyr	Trp	Thr	Glu
				755					760				765
Gly	Lys	Pro	Lys	Ile	Asp	Arg	Ala	Ala	Met	Asp	Gly	Ser	Glu
				770					775				780
Thr	Leu	Val	Pro	Asn	Val	Gly	Arg	Ala	Asn	Gly	Leu	Thr	Ile
				785					790				795
Ala	Lys	Arg	Arg	Leu	Tyr	Trp	Thr	Asp	Leu	Asp	Thr	Asn	Leu
				805					810				815
Ser	Ser	Asn	Met	Leu	Gly	Leu	Asn	Arg	Glu	Val	Ile	Ala	Asp
				820					825				830
Pro	His	Pro	Phe	Gly	Leu	Thr	Gln	Tyr	Gln	Asp	Tyr	Ile	Tyr
				835					840				845
Asp	Trp	Ser	Arg	Arg	Ser	Ile	Glu	Arg	Ala	Asn	Lys	Thr	Ser
				850					855				860
Asn	Arg	Thr	Ile	Ile	Gln	Gly	His	Leu	Asp	Tyr	Val	Met	Asp
				865					870				875
Val	Phe	His	Ser	Ser	Arg	Gln	Ser	Gly	Trp	Asn	Glu	Cys	Ala
				885					890				895
Asn	Gly	His	Cys	Ser	His	Leu	Cys	Leu	Ala	Val	Pro	Val	Gly
				900					905				910
Val	Cys	Gly	Cys	Pro	Ala	His	Tyr	Ser	Leu	Asn	Ala	Asp	Asn
				915					920				925
Cys	Ser	Ala	Pro	Thr	Thr	Phe	Leu	Leu	Phe	Ser	Gln	Lys	Ser
				930					935				940
Asn	Arg	Met	Val	Ile	Asp	Glu	Gln	Gln	Ser	Pro	Asp	Ile	Ile
				945					950				955
Ile	His	Ser	Leu	Arg	Asn	Val	Arg	Ala	Ile	Asp	Tyr	Asp	Pro
				965					970				975
Lys	Gln	Leu	Tyr	Trp	Ile	Asp	Ser	Arg	Gln	Asn	Met	Ile	Arg

-298-

```

      1410              1415              1420
Tyr Val Pro His Pro Ser Ser Leu Ser Gly Ser Leu Pro Gly Met Ser
1425              1430              1435              1440
Arg Gly Lys Ser Met Ile Ser Ser Leu Ser Ile Met Gly Gly Ser Ser
      1445              1450              1455
Gly Pro Pro Tyr Asp Arg Ala His Val Thr Gly Ala Ser Ser Ser Ser
      1460              1465              1470
Ser Ser Ser Thr Lys Gly Thr Tyr Phe Pro Ala Ile Leu Asn Pro Pro
      1475              1480              1485
Pro Ser Pro Ala Thr Glu Arg Ser His Tyr Thr Met Glu Phe Gly Tyr
      1490              1495              1500
Ser Ser Asn Ser Pro Ser Thr His Arg Ser Tyr Ser Tyr Arg Pro Tyr
1505              1510              1515              1520
Ser Tyr Arg His Phe Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val
      1525              1530              1535
Cys Asp Ser Asp Tyr Ala Pro Ser Arg Arg Met Thr Ser Val Ala Thr
      1540              1545              1550
Ala Lys Gly Tyr Thr Ser Asp Leu Asn Tyr Asp Ser Glu Pro Val Pro
      1555              1560              1565
Pro Pro Pro Thr Pro Arg Ser Gln Tyr Leu Ser Ala Glu Glu Asn Tyr
      1570              1575              1580
Glu Ser Cys Pro Pro Ser Pro Tyr Thr Glu Arg Ser Tyr Ser His His
1585              1590              1595              1600
Leu Tyr Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
      1605              1610

```

<210> 765

<211> 207

<212> PRT

<213> Artificial Sequence

<220>

<223> cytoplasmic domain and portion of transmembrane
domain

<400> 765

```

Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
 1      5      10      15
Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
      20      25      30
Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
      35      40      45
Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
      50      55      60
Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser
65      70      75      80
Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
      85      90      95
Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
      100      105      110
Ile Pro Ala Thr Val Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
      115      120      125
Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
      130      135      140

```


Ser	Ala	Ser	Arg	Trp	Lys	Ala	Ser	Lys	Tyr	Tyr	Leu	Asp	Leu	Asn	Ser
145					150					155					160
Asp	Ser	Asp	Pro	Tyr	Pro	Pro	Pro	Pro	Thr	Pro	His	Ser	Gln	Tyr	Leu
				165					170					175	
Ser	Ala	Glu	Asp	Ser	Cys	Pro	Pro	Ser	Pro	Ala	Thr	Glu	Arg	Ser	Tyr
			180					185					190		
Phe	His	Leu	Phe	Pro	Pro	Pro	Pro	Ser	Pro	Cys	Thr	Asp	Ser	Ser	
		195					200					205			

THIS PAGE BLANK (USPTO)